

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



## DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

<b>(51) Classification internationale des brevets <sup>6</sup> :</b> <b>C12N 15/12, C07K 14/47, A61K 31/00</b>	<b>A1</b>	<b>(11) Numéro de publication internationale:</b> <b>WO 97/35974</b> <b>(43) Date de publication internationale:</b> 2 octobre 1997 (02.10.97)
<b>(21) Numéro de la demande internationale:</b> PCT/FR97/00543 <b>(22) Date de dépôt international:</b> 26 mars 1997 (26.03.97) <b>(30) Données relatives à la priorité:</b> 96/03750 26 mars 1996 (26.03.96) FR <b>(71) Déposants (pour tous les Etats désignés sauf US):</b> INSTITUT PASTEUR [FR/FR]; 28, rue du Docteur-Roux, F-75724 Paris Cédex 15 (FR). UNIVERSITE PARIS 7 [FR/FR]; 2, place Jussieu, F-75251 Paris Cédex 05 (FR). <b>(72) Inventeurs; et</b> <b>(75) Inventeurs/Déposants (US seulement):</b> PAULIN, Denise [FR/FR]; 7, rue de Montreuil, F-94300 Vincennes (FR). SMALL, Victor [GB/AT]; Schulstrasse 870, A-5411 Oberalm (AT). MOSSLER, Herber [AT/AT]; Weinbergstrasse 23, A-5201 Seekirchen (AT). LI, Zhenlin [CN/FR]; 33, rue Mazarine, F-75006 Paris (FR). MERIC-SKAY, Mathias [FR/FR]; 9, avenue du Général-de-Gaulle, F-94700 Maisons-Alfort (FR). <b>(74) Mandataire:</b> PHELIP, Bruno; Cabinet Harle & Phelip, 21, rue de la Rochefoucauld, F-75009 Paris (FR).		<b>(81) Etats désignés:</b> AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, brevet ARIPO (GH, KE, LS, MW, SD, SZ, UG), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Publiée</b> <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si de telles modifications sont reçues.</i>
<b>(54) Title:</b> SEQUENCES AHEAD OF GENE SM22, VECTORS CONTAINING SAME, AND THERAPEUTICAL USES THEREOF, PARTICULARLY FOR TREATING VASCULAR DISEASES <b>(54) Titre:</b> SEQUENCES EN AMONT DU GENE SM22, VECTEURS LES CONTENANT ET LEURS UTILISATIONS THERAPEUTIQUES, NOTAMMENT DANS LE TRAITEMENT DES MALADIES VASCULAIRES <b>(57) Abstract</b> <p>DNA sequences including a fragment of the sequence ahead of the coding portion of the SM22 protein gene, or a sequence hybridisable therewith under highly stringent conditions, said fragment being capable of inducing specific gene expression in eukaryotic cells, as well as a sequence coding for a protein or RNA of therapeutic interest, are disclosed. Said sequence or a vector containing same may be used for treating coronary diseases, particularly restenosis.</p> <b>(57) Abrégé</b> <p>Séquences d'ADN comprenant: un fragment de la séquence en amont de la partie codante du gène de la protéine SM22, ou d'une séquence hybridant dans des conditions de forte stringence à ladite séquence en amont, ledit fragment étant susceptible d'induire une expression spécifique d'un gène dans des cellules eucaryotes; et une séquence codant pour une protéine, ou un ARN, d'intérêt thérapeutique. Cette séquence, ou un vecteur la contenant, peut être utilisée pour le traitement des maladies coronariennes, en particulier de la resténose.</p>		

### UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Bésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Caméroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakhstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

**SEQUENCES EN AMONT DU GENE SM 22, VECTEURS LES  
CONTENANT ET LEURS UTILISATIONS THERAPEUTIQUES,  
NOTAMMENT  
DANS LE TRAITEMENT DES MALADIES VASCULAIRES**

5

La présente invention est relative à des séquences en amont d'un gène exprimé dans les cellules du muscle lisse, tel que SM 22, ainsi qu'à des vecteurs les contenant.

Elle a en outre pour objet l'utilisation de ces vecteurs et séquences  
10 en thérapeutique, par exemple pour la production de polypeptides actifs localement ou de manière systémique, notamment des polypeptides immunogènes, des polypeptides régulateurs endogènes tels que les cytokines, ou encore pour la production d'ARN d'intérêt thérapeutique, par exemple un ARN anti-sens. De manière particulière, l'invention a  
15 pour objet l'utilisation des vecteurs et séquences nucléotidiques dans le traitement des maladies vasculaires.

L'athérosclérose est une maladie dégénérative des artères associant les lésions de l'artériosclérose et de l'athérome, et combinant ainsi un durcissement des artères et une dégénérescence graisseuse de  
20 leur tunique interne. Cette insuffisance coronarienne est généralement remédiée par une angioplastie coronaire percutanée consistant à introduire dans le réseau artériel coronaire un cathéter muni à son extrémité d'un ballonnet, d'avancer le cathéter jusqu'au niveau de la sténose et de gonfler le ballonnet afin d'écraser la lésion contre la paroi  
25 et de rétablir ainsi un calibre coronaire permettant une perfusion myocardique suffisante.

Cette technique de revascularisation myocardique est très largement utilisée (50 000 interventions par an en France et 500 000 interventions par an aux Etats-Unis) mais est cependant limitée par la  
30 réapparition, dans les mois suivants l'opération, d'une nouvelle lésion au site dilaté, ou resténose, dans environ 30 % des cas.

Les divers traitements proposés pour traiter ce type de lésions sont résumés dans l'article de FELDMAN et STEG : "Perspective de la  
thérapie génique de la resténose" (Médecine/Sciences, synthèse 1996,  
35 12, 47-55).

Deux mécanismes sont responsables des resténoses : l'hyperplasie intimale et le remodelage artériel. Le premier de ces mécanismes, l'hyperplasie intimale est due à la prolifération des cellules musculaires lisses dans l'intima, qui aboutit à la synthèse d'une volumineuse matrice extra cellulaire. Cette activation serait induite par la libération locale de facteurs de croissance, et de cytokines, ainsi que par la suppression de la synthèse de certains peptides endothéliaux antiproliférants. Le second mécanisme, le remodelage artériel, consiste en une réduction du calibre de l'artère sans modification de l'épaisseur de la paroi, et donc sans prolifération de cellules.

Ces auteurs décrivent diverses tentatives destinées à inhiber l'hyperplasie intimale, par thérapie génique, mais dont les résultats se sont révélés peu satisfaisants, pour diverses raisons. L'inefficacité du transfert est une des raisons généralement invoquée pour expliquer ces résultats. Ce problème a été partiellement résolu par l'utilisation de vecteurs adénoviraux, mais d'autres problèmes sont apparus, liés à l'immunogénicité de ces virus, au risque de recombinaison avec des adénovirus sauvages ou au risque de dissémination de ces vecteurs dans la circulation systémique.

Ces auteurs citent en outre l'utilisation de séquences promotrices spécifiques des cellules vasculaires, comme la préproendothéline, ou des complexes adénovirus-ligand, comme constituant des voies de recherche très intéressantes, sans pour autant développer ces aspects.

La demande WO 96/05.321 (Rhône Poulenc Rorer S.A.) décrit aussi des adénovirus recombinants défectifs comportant un gène suicide pour le traitement de la resténose. Le gène suicide peut être placé sous le contrôle d'un promoteur actif dans les cellules musculaires lisses vasculaires. Seul le promoteur de l'actine  $\alpha$  du muscle lisse est cité dans le cadre de ce mode de mise en oeuvre. Aucun exemple d'utilisation d'une telle construction ne vient illustrer cette demande.

Les protéines constituant les muscles lisses ont fait l'objet d'études parcellaires. Ainsi, la protéine SM 22, qui constitue un des premiers marqueurs à apparaître lors de la différenciation des cellules de muscles lisses, a été étudiée et la partie codante de son gène a été séquencée

chez la souris (SOLWAY et AL, 1995, J. Biol. Chem, 270.22, 13460-13469) et chez le rat (KEMP et AL, 1995, BIOCHEM. J. 310,1037-1043).

Les séquences du gène de poulet et du gène humain ont aussi été déterminés (Nishida et al, 1991, Biochem. Int., 23,663-668 pour le gène de poulet ; Thweatt et al, 1992, Biochem. Biophys. Res. Comm., 187,1-7 pour le gène humain).

Le gène codant pour la protéine SM 22 alpha chez la souris est composé de cinq exons et comprend 6,2 kb. Il est présent en une seule copie dans le génome. SOLWAY et al. ont montré que les 441 paires de base en amont de la région codante du gène étaient nécessaires et suffisantes pour induire un taux important de transcription d'un gène reporteur, dans des cultures cellulaires primaires d'aortes de rats et dans les lignées A7r5. Ils ont d'autre part mis en évidence que les ARN messagers de ce gène étaient exprimés à des taux importants dans l'aorte, dans l'utérus, dans les poumons et dans l'intestin.

KEMP et al (1995, BIOCHEM. J. 310, 1037-1043.) ont cloné et séquencé le fragment de 1,9 kb en amont de la séquence codante du gène SM 22 de rat. Des délétions effectuées dans le fragment ont montré que la région du promoteur comprise entre les nucléotides +65 et -303 étaient plus actives qu'un fragment de 1,5 kb comprenant une partie importante de la région du gène non traduite, ce qui suggérerait la présence de séquences régulatrices à l'extrémité 5' du promoteur.

OSBOURN et AL (1995, Gene, 154, 249-253) ont aussi étudié la région 5' en amont du gène SM 22 du rat. Ils ont mis en évidence la présence de deux introns dans cette région.

On notera que ces études ont été faites in vitro en utilisant des cultures de cellules primaires de l'aorte, et non in vivo. Or, ces conditions d'expérimentation ne sont pas représentatives de l'activité des gènes dans les cellules de l'organisme.

De plus, elles ne concernent qu'une région très limitée de la partie non codante du gène située à l'extrémité 5'. c'est-à-dire la partie en aval du nucléotide -441.

Plus récemment Li et al. (1996, J. Cell. Biol, 132, 849-859) ont étudié la régulation du gène SM22 à l'aide de souris transgéniques.

Néanmoins, cet article ne mentionne aucune application de vecteurs portant des fragments de ce gène, et en tout état de cause aucune application thérapeutique.

En outre un nombre limité de constructions ont été fabriquées, et de manière générale seul l'effet sur des embryons, et non sur des adultes, a été testé.

Les auteurs déduisent de leur étude que les séquences comprises entre les nucléotides -2735 et -445 ne contiennent pas d'éléments régulateurs musculaires essentiels.

10 Le demandeur s'est attaché à identifier in vivo, sur des individus adultes l'activité des diverses régions constituant la région en amont de la séquence codante exprimée du gène de la protéine SM 22. Il s'est d'autre part attaché à déterminer si des modifications dans cette séquence induisait les changements dans l'expression du gène dans les  
15 différents tissus.

Il a mis en évidence de manière surprenante qu'il était possible, en modifiant la région en 5' du gène de la protéine SM 22 de souris, d'obtenir, chez des individus adultes, une expression spécifique d'un gène reporteur dans les muscles lisses des artères et en particulier de  
20 l'aorte.

Il a en outre montré que les séquences identifiées par LI et al. (1996, précédemment cités) comme ne contenant pas de séquences régulatrices essentielles, contenaient au contraire des séquences indispensables à une expression spécifique du gène SM 22 chez  
25 l'adulte.

Il a enfin montré de manière surprenante que des séquences introniques étaient nécessaires pour l'induction d'une expression spécifique dans les artères.

La présente invention a donc pour objet une séquence d'ADN  
30 caractérisée en ce qu'elle comprend :

- un fragment de la séquence en amont de la partie codante du gène de la protéine SM 22, ou d'une séquence hybridant dans des conditions fortement stringentes à ladite séquence en amont, ledit fragment étant susceptible d'induire une expression spécifique in vivo de  
35 ce gène dans des cellules des artères, et

- une séquence codant pour une protéine, ou un ARN, d'intérêt thérapeutique .

L'invention concerne également toute séquence nucléotidique ou oligonucléotide d'origine naturelle ou obtenue par synthèse chimique  
5 présentant une homologie avec la séquence SEQID n°1 d'au moins 70 %.

On entend par séquence nucléotidique d'origine naturelle, un fragment d'ADN complémentaire obtenu par transcription inverse de l'ARN messager cellulaire ou encore un fragment d'ADN génomique  
10 obtenu après clivage de l'ADN cellulaire à l'aide d'enzymes de restriction.

On entend par séquence nucléotidique obtenue par synthèse chimique un fragment d'ADN de séquence connue généré par synthèse automatique de polynucléotides, par exemple à l'aide d'un appareil  
15 automatique adapté.

On entend par protéine ou ARN d'intérêt thérapeutique une molécule susceptible d'inhiber la croissance des cellules du muscle lisse, d'activer la croissance des cellules endothéliales, de consolider les artères et/ou capable d'induire une réponse immunitaire cellulaire ou  
20 humorale.

Pour la présente invention, le terme "conditions fortement stringentes" est utilisé dans le sens donné par Maniatis et al., 1982 (Molecular cloning , A Laboratory Manual, Cold Spring Harbor Lab. CSH, N.Y. USA ou l'une de ses récentes rééditions. A titre préféré, les  
25 conditions d'hybridation sont telles que décrites par Maniatis (édition de 1982). Trois lavages destinés à éliminer les fragments non hybridés sont nécessaires et sont effectués à 65°C en présence de 0,2 SSC et 0,1 % SDS, en l'absence de formamide.

De manière avantageuse, ladite séquence comprend un fragment  
30 de la séquence comprise entre les nucléotides -2126 et +4135, dont la séquence SEQ ID n°1 est donnée en annexe, de la séquence comprise entre les nucléotides -2126 et +65 dont la séquence SEQ ID n° 2 est donnée en annexe ou encore de la séquence comprise entre les nucléotides -2126 et -445  
35 dont la séquence SEQ ID N°3 est donnée en annexe.



Elle peut aussi comprendre un fragment de la séquence SEQ ID N° 4.

Un tel fragment peut en particulier comprendre au moins une partie de la séquence SEQ ID N°5 qui correspond à la partie en amont de l'extrémité 5' de la partie codante comprise entre les nucléotides -2126 et -1340. Une telle séquence SEQ ID N°5 dont la séquence est donnée en annexe constitue un objet de la présente invention.

Les séquences SEQ ID n° 2 et SEQ ID n° 1 sont comprises respectivement dans les plasmides p2126nlz et p2126INTnlz portées par les souches déposées le 25 mars 1996 auprès de la Collection Nationale de Culture de Microorganismes de l'Institut Pasteur (CNCM) respectivement sous les n° I-1685 et I-1686.

L'invention a également pour objet les séquences hybridant dans des conditions de forte stringence avec les séquences SEQ ID n° 1, SEQ ID n° 2, SEQ ID n° 3 et SEQ ID N° 5.

Les séquences incluant la séquence du premier intron du gène SM 22, et en particulier la séquence SEQ ID N° 1 constituent des modes de mise en oeuvre particulièrement avantageux de la présente invention, car elles permettent une expression de protéines ou ARN d'intérêt thérapeutique, chez l'adulte, spécifique des artères.

La présente invention n'est néanmoins pas restreinte à des séquences contenant des fragments du gène de souris, mais concerne toute autre séquence de toute autre espèce ayant les mêmes propriétés, à savoir d'induire spécifiquement une expression d'un gène dans les cellules des artères. Ainsi, l'homme du métier pourra avantageusement se référer à l'article de KEMP et AL (1995, précédemment cité) qui décrit la séquence en amont du gène de rat de la protéine SM 22. La présente invention couvre aussi toute séquence comprenant des fragments des séquences en amont de gènes de la protéine SM 22, modifiées par exemple par délétion de certaines structures qui conservent des fonctions identiques ou similaires à celles de la séquence complète.

La protéine d'intérêt thérapeutique peut être une protéine induisant la formation d'un composé cytotoxique, telle que la thymidine kinase du virus de l'herpès. Cette protéine présente la particularité de transformer le ganciclovir (Merck Index, référence 4262), un analogue de la

guanosine, en un dérivé qui bloque la synthèse de l'ADN dans les cellules en réplication. L'introduction d'un gène codant pour cette protéine dans les séquences selon la présente invention permet donc de bloquer la prolifération des cellules dans le cas de l'hyperplasie intima-

5 Une séquence codant la thymidine kinase est notamment décrite par Caruso et Klatzmann (1992, Proc. Natl. Acad. Sci., USA, 89, 182-186).

Ladite protéine d'intérêt thérapeutique peut aussi être une protéine présentant un effet cytostatique, telle que la protéine codée par le gène Rb (Chang et al., 1995, Science, 267, 518-522), ou par le gène eNOS  
10 (Hamon et al., 1994, Circulation, 90, 1357-1362)..

Elle peut aussi être une protéine présentant une activité lipolytique, telle qu'une lipoprotéine lipase. De telles protéines sont codées par les séquences décrites par Reyner (1995, Nature Genetics, 10, 28) et Ming Sun Liu (1994, J. Biol. Chem, 269-11417).

15 Elle peut aussi être un facteur de croissance des cellules endothéliales. Une telle protéine peut être en particulier une interleukine.

Elle peut être une protéine musculaire, telle que la myosine, ou l'actine, ou encore une protéine de structure tissulaire telle que le  
20 collagène ou l'élastine.

Elle peut être une protéine induisant une réponse immunitaire telle que décrite dans la demande WO 90 11092, et en particulier un antigène viral tel que la glycoprotéine gp120 ou la protéine Nef du VIH (virus de l'immunodéficience humaine).

25 De manière générale, ladite protéine peut être une ou plusieurs protéines présentant un intérêt thérapeutique ou vaccinal, telles que des protéines d'intérêt immunothérapeutique, par exemple des interleukines, les facteurs de croissance, par exemple les facteurs de croissance des fibroblastes (FGF) ou le NGF (Nerve Growth Factor), des protéines  
30 pouvant induire une réponse immunitaire, que ce soit une immunité humorale, cytotoxique ou cellulaire, ou des protéines permettant de compléter l'activité d'un gène normalement exprimé chez l'individu à traiter mais qui ne l'est plus, soit par mutation ou délétion de sa séquence.

Cette séquence peut aussi coder pour un ARN d'intérêt thérapeutique. Un tel ARN peut être l'ARN antisens de la protéine p 53 ou un ARN tel que décrit dans la demande WO 90 1 1092. déposée par la société VICAL.

5        La présente invention a en outre pour objet des vecteurs caractérisés en ce qu'ils contiennent une séquence selon l'invention, telle que décrite ci-dessus. Un tel vecteur peut contenir tout autre séquence d'ADN nécessaire à l'expression de la protéine, ou de l'ARN, d'intérêt thérapeutique dans les tissus cibles, et en particulier peut  
10        contenir une origine de réplication efficace dans les cellules des artères, en particulier dans les cellules du muscle lisse.

Un tel vecteur peut comprendre des séquences permettant la recombinaison homologue chez l'organisme traité, spécifique du gène à remplacer, lesdites séquences étant placées en amont et en aval de la  
15        séquence selon l'invention. Du fait de la présence de telles séquences, le gène non désiré, présent dans l'organisme traité sera remplacé par le gène porté par le vecteur ou la séquence et que l'on désire voir s'exprimer dans l'organisme.

Une telle méthode de recombinaison homologue peut être du type  
20        de celle décrite par Le Mouellic et al., (1990, Proc. Nat. Acad. Sci. USA, 87, 4712-4716) ou dans la demande PCT WO 91/06.667.

Un tel vecteur peut être un vecteur dérivé d'un adénovirus. Des adénovirus adaptés à la mise en oeuvre de la présente invention sont en particulier ceux décrits par FELDMAN et STEG (1996, précédemment  
25        cité) ou OHNO et al (1994, Sciences, 265, 781-784) ou encore dans la demande FR 94.03.151 (Institut Pasteur, Inserm). Ces virus sont généralement rejetés par l'individu, du fait de leur trop forte immunogénicité. Néanmoins, des souches de ces virus ont été sélectionnées afin de diminuer leur caractère immunogène. En tout état  
30        de cause, ces vecteurs, même présentant une forte immunogénicité, peuvent être utilisés dans le cadre de la présente invention, car les artères sont traitées durant des périodes de temps relativement courtes, de l'ordre de quelques semaines, compatibles avec les temps de rejet des adénovirus par le système immunitaire de l'hôte.

On notera néanmoins que l'introduction des séquences selon la présente invention dans les vecteurs décrits ci-dessus n'est pas indispensable, et que les cellules des artères peuvent être directement transfectées par de l'ADN comprenant ces séquences.

5 Les séquences nucléiques selon la présente invention peuvent être introduites après couplage covalent de l'acide nucléique avec des composés favorisant leur pénétration dans les cellules ou leur transport vers le noyau, les conjugués résultants étant éventuellement encapsulés dans des microparticules polymères, comme dans la demande  
10 internationale WO 94/27238 de Meidsorb Technologies International.

Selon un autre mode de réalisation, les séquences nucléiques peuvent être incluses dans un système de transfection comprenant des polypeptides favorisant leur pénétration dans les cellules, comme dans la demande internationale WO 95/10534 de Seikagaku Corporation.

15 Ces vecteurs ou séquences peuvent être administrés in situ par tout moyen connu de l'homme du métier. Ainsi, ils peuvent être délivrés in situ à l'aide d'un cathéter à ballonnet recouvert de canaux, tel que décrit par FELDMAN et STEG (1996, précédemment cité).

Ils peuvent aussi être administrés au cours d'une opération, ou en  
20 dénudant l'aorte. Un autre mode d'administration consiste à introduire un grillage de faible maillage imprégné d'ADN comprenant ces séquences, le long de l'aorte ; tel que décrit par Feldman et al. (1995, J. Clin. Invest., 95, 2662-2671).

Ces séquences ou vecteurs peuvent avantageusement être  
25 administrés sous la forme d'une composition les contenant, par exemple dans un gel facilitant leur transfection dans les cellules des artères. Un tel gel peut être un complexe de poly-L lysine et de lactose, tel que décrit par MIDOUX (1993, Nucleic Acid Research, 21, n°4, 871-878) ou du poloxamer 407 tel que décrit par PASTORE (1994, Circulation, 90, 1-517). Ils peuvent aussi être mis en solution dans une solution  
30 tamponnée ou être associés à des liposomes.

La présente invention est en outre relative à des médicaments contenant de telles séquences, vecteurs ou compositions, ainsi qu'à des compositions pharmaceutiques les contenant en quantités

pharmaceutiquement efficaces ainsi que des excipients pharmaceutiquement compatibles.

De tels séquences, vecteurs ou compositions peuvent être avantageusement utilisés pour la fabrication de médicaments pour la  
5 délivrance au niveau des artères d'un ADN (ADNc ou ADN génomique) pouvant exprimer notamment un gène codant pour une protéine d'intérêt, pour le traitement des maladies coronariennes, et en particulier pour le traitement de la resténose. Elles peuvent néanmoins être aussi  
10 utilisées pour la fabrication d'un médicament pour le traitement de mutations fragilisant les vaisseaux. Un tel médicament pourrait, par exemple, comprendre une séquence ou un vecteur selon l'invention capable d'exprimer un gène codant pour une protéine normale de la desmine ou une protéine d'adhésion du muscle du type CAM.

La présente invention a aussi pour objet des ARN exprimés à partir  
15 des séquences et vecteurs objets de la présente invention, et en particulier des ARN messager.

La présente invention a aussi pour objet un procédé de criblage de molécules in vitro, en vue de tester leur activité sur les séquences régulatrices du gène codant pour la protéine SM 22. Un tel procédé  
20 comprendra, par exemple, une première étape suivant laquelle on transfecte des cellules avec une séquence ou un vecteur selon l'invention, comprenant un gène reporteur placé sous le contrôle de tout ou partie de la séquence du promoteur du gène codant pour la protéine SM 22. Dans une seconde étape, les cellules ainsi transfectées sont  
25 incubées en présence de la molécule à tester, puis l'expression du gène reporteur est quantifiée.

Les cellules utilisées pour la transfection sont avantageusement des cellules de muscle lisse, en particulier des cellules de l'aorte, soit sous la forme d'une culture primaire, soit sous la forme d'une lignée  
30 cellulaire stable.

Dans le cas où les cellules sont transfectées par des constructions telles que p2126nlz ou p2126INTnlz, la quantification de la bêta-galactosidase produite sera avantageusement effectuée par lecture  
35 optique, par exemple en utilisant la trousse de détection commercialisée par Boehringer sous la référence 1 669 893 (catalogue Boehringer-

Mannheim - « Biochemicals », 1996, page 236, bêta-Gal Reporter Gene Assay, Chemiluminescent).

Dans un autre mode de réalisation du procédé de criblage in vitro selon l'invention, le gène reporteur est le gène codant pour la luciférase.

5 Dans ce cas, la quantification de l'expression de la luciférase est avantageusement effectuée par chimioluminescence, par exemple au moyen de la trousse de diagnostic commercialisée par Boehringer sous la référence 1 758 241 (catalogue Boehringer-Mannheim « Biochemicals », 1996, Luciferase Reporter Gene Assay).

10 La présente invention est aussi relative à des animaux transgéniques et en particulier des souris portant une séquence ou un vecteur tel que défini ci-dessus dans lequel le gène codant pour la protéine d'intérêt thérapeutique est remplacée par un gène reporteur.

La présente invention est enfin relative à un procédé de détection  
15 de mutations de la séquence en amont de la partie codante du gène de la protéine SM22. De telles mutations sont en effet susceptibles de modifier l'expression du gène codant pour la protéine SM22, en particulier diminuer ou même abolir l'expression de ce gène dans les muscles lisses. De telles mutations seraient de nature à entraîner une  
20 modification du fonctionnement du muscle lisse, la protéine SM22 étant apparentée à une famille de protéines impliquées dans la régulation de la fixation du calcium, du type calponine (Ayme-Southgate et al, 1989, J. Cell. Biol.). Un tel procédé de détection de mutations peut être avantageusement le procédé objet de la demande de brevet français  
25 FR-93 10821.

De manière complémentaire, des séquences selon l'invention modifiées de telle sorte qu'elles permettent une augmentation de l'expression du gène SM22 font également partie de l'invention.

De telles souris transgéniques peuvent être utilisées pour cribler  
30 des molécules pour leur activité sur les séquences régulatrices gène codant pour la protéine SM 22. Des molécules peuvent être administrées à des souris, puis après sacrifice, des coupes histologiques sont effectuées afin de mettre en évidence les tissus colorés par le gène reporteur.

Pour la mise en oeuvre de la présente invention, l'homme du métier pourra avantageusement se référer au manuel suivant : "SAMBROK et al (Molecular cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1989), ou à l'une de ses récentes rééditions.

La présente invention est illustrée sans pour autant être limitée par les exemples qui suivent :

La figure 1 représente la carte de restriction et la structure exon/intron du site SM 22, comprenant la région -2474 à +6030 (par rapport au site d'initiation de la transcription). Les exons sont représentés par des carrés blancs tandis que les flèches indiquent la position des séquences répétées de type B1.

La figure 2 représente l'alignement de trois séquences répétées de type B1, dans lesquelles les échanges de base sont indiquées.

La figure 3 est un autoradiogramme illustrant l'identification du site d'initiation de la transcription. Le puits P correspond à l'amorce utilisée.

La figure 4A représente la séquence de la région 5' en amont du site d'initiation de la transcription, et de l'exon du gène SM 22. Les sites de fixation des différents facteurs sont indiqués, ainsi que les différentes séquences consensus. Les lignes discontinues indiquent la position des séquences répétées de type B1.

La figure 4B est une carte de restriction schématique du gène SM 22.

La figure 4C est une carte de restriction détaillée du gène SM 22.

La figure 5 illustre la fabrication du plasmide p352nlz.

La figure 6 illustre la fabrication du plasmide p2126 nlz.

La figure 7 illustre la fabrication du plasmide p2126INTnlz.

Les figures 8A et 8B illustrent l'expression du gène lacZ dans des embryons de souris transgéniques. Les figures 8A, 8B, représentent respectivement des embryons aux stades 12,5 et 15,5 jours.

Les figures 9A à 9D représentent des coupes histologiques d'embryons colorés à l'aide de lacZ. La figure 9A représente une coupe d'un embryon au stade 12,5 jours au niveau du coeur, montrant une coloration intense exclusivement dans le myocarde du ventricule droit (RV) et dans les parois des artères. (abréviations : AA : coude de la

quatrième artère aortique ; CA : artère carotidienne ; E : oesophage ; lv : ventricule gauche ; PA : partie proximale de l'aorte ; PT : tronc pulmonaire ; RA : atrium droit ; RV : ventricule droit ; T : trachée ; V : veine cardinale droite et gauche).

5 La figure 9B représente une coupe à travers la région abdominale d'un embryon au stade 14,5 jours, montrant une coloration des artères ombilicales (UA) et l'absence de marquage des viscères (B: vessie ; D : duodenum ; H : gros intestin ; L: foie ; M : intestin grêle ; ST : estomac).

La figure 9D est une coupe à travers les segments de la queue au  
10 stade 12,5 jours montrant le marquage dans le myotome (my) et l'artère de la queue (a).

La figure 9C est un agrandissement de la partie de la photographie correspondant à l'artère ombilicale au stade de 12,5 jours, montrant un marquage exclusif de la couche musculaire (m) et une absence de  
15 marquage de l'endothélium (e).

Les figures 10A à 10C illustrent l'expression du transgène dans des souris adultes.

La figure 10A représente une vue de dessus de l'ensemble du coeur montrant un intense marquage de l'aorte (a) et de l'artère  
20 pulmonaire (pa) et une absence de marquage de la veine cave (vc). Les figures 10B et 10C sont des coupes à travers la couche de muscle lisse (sm) du colon de l'adulte et d'une artère mésentère adjacente (ma). Le marquage de lacZ (figure 10B) montre que l'expression du transgène est restreinte à l'artère, tandis que l'immunofluorescence avec des anticorps  
25 SM 22 (figure 10C) met en évidence l'expression du gène SM 22 endogène dans les cellules du muscle lisse aussi bien de l'artère que du colon.

La figure 11 représente un vecteur comportant les régions régulatrices du gène SM 22 et le gène codant la thymidine kinase du  
30 virus de l'herpès.

Les figures 12A à 12D illustrent l'expression de la construction SM 445 nlz dans des embryons de 15,5 jours p.c. (figures 12A à 12C) et dans un adulte de 2 mois (figure 12D).



**EXEMPLES:****Matériels et méthodes****1. Clonage et caractérisation du gène de la souris SM 22**

Environ  $10^6$  phages recombinants d'une banque génomique de  
souris c57/bl construite dans le phage delta-EMBL 3 (fourni par Dr D.  
Plachov, Münster, Allemagne) ont été criblées en utilisant le fragment  
BAL I/Eco RV de 890 bp de l'ADN complémentaire de la protéine de  
souris SM22 ( Almendral et al, 1989 ; Exp. Cell Res. 181, 518-530),  
comme sonde. La sonde a été marquée radioactivement par initiation  
aléatoire et l'ADN a été hybridée sur des filtres laissés durant une nuit à  
65°C dans un tampon de Church. Les filtres ont été ensuite lavés dans  
du 0,2 X SSC/0.1 % SDS à 65°C et des plaques positives ont été  
purifiées et homogénéisées par trois recriblages successifs dans des  
conditions identiques. L'un des clones isolés s'est révélé contenir un  
locus SM 22 entier. Il a été ensuite cartographié à l'aide d'enzymes de  
restriction et sous-cloné. Afin d'éviter des réarrangements durant les  
procédures de clonage, une analyse par hybridation de type Southern  
de l'ADN génomique de souris et des clones génomiques a été effectuée  
et les figures de restriction des fragments ont été comparées.

La séquence du gène SM 22 a été déterminée sur les deux brins  
par la méthode de terminaison de chaîne ( Sanger et al, 1977, Proc.  
Natl. Acad. Sci. 74, 5463) en utilisant le kit Sequenase V2.0 (United  
States Biochemical, Cleveland, Ohio).

**2. Hybridation de type Northern, extension par amorce,  
5'RACE, et analyse de protection par RNase**

De l'ARN total de lignées cellulaires et de tissu de souris adulte a  
été isolée par la méthode de Chirgwin et al, (1979, Biochemistry, 18.  
5294-5299) modifiés. Pour l'isolement des ARN messager, le kit Oligotex  
(Quiagen Inc, Chatsworth, CA) a été utilisé en suivant les instructions du  
fabricant.

Les analyses par hybridation de type Northern ont été effectuées  
en utilisant des gels d'agarose contenant du formaldéhyde et par  
hybridation capillaire sur des membranes Hybond-N (Amersham Life  
Science, Little Chalfont, Royaume Uni) selon les méthodes connus de  
l'homme du métier.

Pour les analyses par extension d'amorce, un oligonucléotide synthétique de 30 nucléotides (P27N1) complémentaire de la séquence +36 à +65 de l'ADN complémentaire du gène SM 22 présentant la séquence SEQ ID N° 6 :

5 (5'-AAGGCTTGGTCGTTTGTGGACTGGAAGGAG-3), a été marqué par de la polynucléotide kinase T4. 5 µg d'ARN messenger d'utérus de souris ont été hybridés à 55°C avec  $1-2 \times 10^5$  cpm de sonde purifiée et la réaction d'extension d'amorce a été effectuée en utilisant les procédures de laboratoire connues de l'homme du métier. Les  
10 résultats sont analysés par électrophorèse sur des gels de polyacrylamide dénaturants à 6 %. Une réaction utilisant la même amorce a été chargée dans un puits adjacent afin de permettre une détermination directe du site d'initiation de la transcription.

La méthode 5'RACE a été effectuée en utilisant deux  
15 oligonucléotides synthétiques antisens de 18 bp correspondant aux bases +112 à +129 (amorce 2) et +162 à +179 (amorce 1) de l'ADN complémentaire SM 22, tel que décrit en substance par FROHMAN et al, (1988, Proc. Natl. Acad. Sci, 85, 8998-9002).

En vue des analyses de protection par la RNase, un fragment  
20 obtenu par PCR de 417 bp comprenant la région -352 à +65 du gène du SM 22 a été fabriqué et cloné par la méthode des bouts collants dans le site *HinC II* du vecteur pBluescript SK+. La séquence du fragment obtenu par PCR a été vérifiée. Afin de générer une sonde ARN antisens marquée radioactivement, la construction a été linéarisée avec *Sph I* à  
25 la position -203 du gène SM 22 et transcrite avec de la RNA polymérase T7 en présence de  $^{32}\text{P}$ -CTP. La sonde ( $3 \times 10^5$  cpm) a été hybridée avec 5 µg d'ARN messenger d'utérus de souris à 42°C durant une nuit. Après hybridation, les échantillons ont été incubés avec des RNases A et T (1 heure à 37°C), précipités, et analysés comme décrits pour  
30 l'analyse par extension d'amorce.

### **3. Culture cellulaire et transfection**

Les lignées cellulaires NIH 3T3, 10T1/2 et 8/47 ont été cultivées dans du DMEM (Life Technologies, Inc, Vienne, Autriche) complémentées avec 10 % de sérum de fœtus de veau (Hyclone  
35 Inc., Etats-Unis) à 37°C, dans une atmosphère de  $\text{CO}_2$  à 7%.

Pour les expériences de transfection, des cellules confluentes ont été étalées, repiquées dans des boîtes de Petri de 6 cm, et laissées croître jusqu'à obtenir approximativement une confluence de 75 % avant ajout du mélange de transfection.

5 Les transfections ont été effectuées en utilisant le réactif Lipofectamine commercialisé par Life Technologies Inc..

En résumé 24 µg de Lipofectamine ont été ajoutés à 6 µg de plasmide portant le gène reporteur superenroulé et 2 µg de PCMVβ gal comme témoin interne. Après ajout de 3 ml de DMEM, le cocktail de  
10 transfection a été ajouté aux cellules. Au bout de 6 heures, ce mélange a été remplacé avec du milieu de croissance contenant 10 % de sérum et à nouveau changé au bout de 12 heures. L'expression du produit du gène reporteur a été analysée après 24 heures. Pour les expériences de stimulation par le sérum, des cellules transfectées ont été placées dans  
15 du DMEM contenant 20 % de sérum, 24 heures avant d'être récoltées.

#### **4. Dosage de la CAT et de la β-galactosidase**

Les cellules ont été prélevées sur des boîtes de pétri avec un racloir à cellules, centrifugées, resuspendues dans 150 µl de tampon Z (100 mM de phosphate de sodium, pH 7,5, 1mM MgCl<sub>2</sub>, 10 nM KCl, 50  
20 nMβ - mercaptoethanol) et lysées par trois cycles de congélation/décongélation. Les débris ont été éliminés par centrifugation à basse vitesse et l'activité de la β-galactosidase a été déterminée à partir de 2 à 20 µl de surnageant selon les procédés connus de l'homme du métier. Pour la détermination de l'activité CAT  
25 selon GORMAN et al (1982 Mol. Cell. Biol., 2, 1044-1051), de 10 à 100 µl de lysat ont été utilisés. Afin de tenir compte des variations dans l'efficacité de transfection, l'activité CAT a été normalisée avec l'activité βgal.

#### **5. Production des souris transgéniques**

30 Des inserts des plasmides pnlz2126 et p2126INTnlz, déposés auprès de la CNCM respectivement sous les n°I-1685 et n°I-1686, ont été isolés par digestion avec Sal I/Pme I, suivie d'une électrophorèse sur gel, puis d'une purification utilisant le kit Geneclean (Bio 101, la Jolla, CA) puis resuspendus dans 10 mM de Tris Hcl pH 7,5, 0,25mM EDTA.

Les animaux transgéniques ont été produits comme décrits précédemment par LI et al (1993, Development, 177/3, 947-959).

Des souris portant le transgène ont été identifiées par PCR et hybridation de type Southern en utilisant une sonde ADN. La  
5 construction pnlz2126 a permis d'obtenir trois animaux positifs sur 17 souris, deux d'entre elles transmettant et exprimant cette construction.

Pour la construction p2126INTnlz deux animaux positifs sur les 28 testés ont été obtenus et à nouveau un seul exprime et transmet cette construction.

10 Les souris transgéniques ont été rétro-croisées avec des souris c57/bl et les colorations histochimiques ont été effectuées avec des souris hémizygotes pour le transgène.

#### **6. Colorations histochimiques**

Des embryons entiers ou des tissus d'adultes ont été fixés durant  
15 15 à 30 minutes à 4°C dans du formaldéhyde à 1 % dans du tampon A (100 mM de phosphate de sodium , pH 7,3, 2 mM MgCl<sub>2</sub>, 0,1% de sodium de deoxycholate, et 0,2 % NP-40). Après rinçage, les spécimens ont été colorés durant une nuit à 30 °C dans du tampon A contenant 1 mg/ml de X - gal (Sigma, St Louis, Missouri), 5 mM de ferrocyanure de potassium,  
20 5mM de ferricyanure de potassium, et 20 mM de Tris-Cl pH 7,3. Les échantillons colorés ont alors été rincés, déshydratés à l'aide de concentrations croissantes d'éthanol, clarifiés dans du xylène, et enrobés dans de la paraffine. Les coupes histologiques ont été effectuées avec une épaisseur de 7 à 10 µm et à nouveau colorées avec  
25 de l'hématoxiline d'Ehrlich et de l'éosine. Les embryons ont d'autre part pu être à nouveau clarifiés à l'aide d'un mélange d'alcool benzylique et de benzoate de benzyle.

Pour l'immunohistocoloration, les tissus de souris ont été fixés dans 3 % de paraformaldéhyde dans du tampon PBS à 4°C durant 3  
30 heures, puis enrobés dans un milieu Tissue-Tek (Reichert-Jung, Vienne) et sectionnés sur un cryo-microtome à des épaisseurs de 5 à 10 µm. Les coupes sélectionnées ont été montées et colorées en utilisant un anticorps monoclonal dirigé à l'encontre de la SM 22 (Duband et al, 1993, Differentiations, 55, 1-11), à une dilution 1/30 ainsi qu'en utilisant  
35 un anticorps secondaire marqué Cy3 (Biological Detection Systems,

USA). Les coupes ont été examinées à l'aide d'un microscope à fluorescence.

### **EXEMPLE 1.**

#### **Caractérisation du locus SM 22**

5 Le criblage par hybridation de la banque génomique de souris a conduit à l'isolement de trois clones se recouvrant, contenant le gène complet de la SM 22. Une carte de restriction, qui est partiellement reproduite sur la figure 1 a été établie et les exons ont été localisés par hybridation à l'aide de sondes oligonucléotidiques synthétiques.

10 Les régions présentant un intérêt ont été sous clonées et séquencées.

La séquence a été déterminée entre les positions -2474 et +110 en aval du site de polyadénylation et a été comparée dans la banque de données EMBL. Le gène couvre 5923 bp à partir du site de début de la transcription jusqu'au site de polyadénylation et la séquence codante est  
15 partagée en 5 exons (figure 1). Toutes les frontières exon-intron correspondent à des séquences consensus.

Le premier exon contient seulement une séquence leader en 5' et le codon de départ de la transcription est donc localisé dans le second  
20 exon séparé du site de départ de la transcription par plus de 4kbp de séquence d'intron.

Deux signaux de polyadénylation potentiels ont été localisés au position 5905 et 5913 et plusieurs régions TGT placées en aval du site de polyadénylation, prennent vraisemblablement part dans la  
25 terminaison de la transcription.

L'analyse de la séquence révèle de plus la présence de trois séquences répétées (figure 2) qui présentent des homologies importantes avec les éléments répétés de type B1, dont certains sont transcrits par la RNA polymérase III et codent pour des ARN 5S. On  
30 notera que toutes les séquences répétées sont dans une orientation inverse par rapport au locus SM 22.

Le site de départ de la transcription a été cartographiée par analyse par extension d'amorce (figure 3). Un oligonucléotide antisens complémentaire de l'extrémité 3' de l'exon 1 conduit à l'obtention de  
35 quatre produits d'extension qui diffèrent seulement par une paire de

bases (bp) en longueur, le plus long d'entre eux étant de 65 bp. Le produit de la transcription démarre donc avec un G, 77bp en amont du codon d'initiation de la transcription.

Les produits d'extension les plus courts ont été vraisemblablement  
5 causés par une terminaison prématurée de l'extension de l'amorce due à des modifications (capping) à extrémité 5' de l'ARN messager.

Ces résultats ont été confirmés par l'analyse par protection par la RNase, qui conduit à un fragment protégé de 65bp et par le clonage et le séquençage du produit d'extension dérivé d'une réaction  
10 indépendante avec une amorce antisens située dans le second exon du gène.

Ces résultats sont en accord avec ceux de Solway et al (1995 précédemment cité). Des différences de séquences mineures entre ces deux résultats sont vraisemblablement dues à des variations alléliques  
15 entre les différentes souches de souris.

La séquence TTAAA en position -28bp (figure 4) est en relation étroite avec la séquence TATAAA et a vraisemblablement la fonction d'une boîte TATA. L'analyse par ordinateur de la séquence en 5' révèle la présence de plusieurs sites potentiels de liaison pour des facteurs  
20 ainsi que pour des éléments qui sont connus comme contribuant à la régulation de la transcription de gène musculaires (figure 4). Au total 11 boîtes de type E (CANNTG), quatre motifs du type Mef-2/rSRF (YTAWAAATAR), quatre éléments de liaison SRF potentiels (CC(AT)6GG), cinq sites de liaison AP-2 (CCCMNSSS) et cinq motifs  
25 SP1 (GGGCGG) ont été localisés. Enfin, un élément similaire au motif TGT3-3, récemment identifié comme site de liaison d'un facteur nucléaire du muscle lisse et comme contribuant à la régulation de la transcription de l'actine alpha du muscle lisse, a été localisé (figure 4).

## **EXEMPLE 2**

### **30 Clonage et construction de plasmides recombinants pour la création de souris transgéniques**

#### **1. Fabrication de pnlz**

Le plasmide pGEM7 (c'est-à-dire le plasmide pGEM-72F (+/-), commercialisé par Promega Corp., Madison, Wi, USA. Séquences  
35 disponibles dans Gen Bank, sous les numéros suivants : X65310 et X

65311) comprend un gène de résistance à l'ampicilline, un gène lac Z et une origine de réplication. Il comprend 3000 paires de base. Il a été digéré avec Nae I/Bsp 120I. L'extrémité sortante Bsp 120I du vecteur a été rendue « bouts francs » avec l'enzyme de Klenow, et le plasmide a été refermé en présence d'un linker Sal I de 10 pb. Le fragment lacZ de pGEM 7 a donc été délété, un site Sal I introduit, et un site Bsp 120I restauré.

Un site Pme I a été introduit par insertion d'un linker Pme I avec des extrémités simple-brin sortantes Nsi I, dans le site Nsi I de la construction.

Finalement, le gène lacZ couplé au signal de localisation nucléaire a été isolé à partir de pDes2.nlz (LI et al. 1993, précédemment cité) avec Hind III et Sac I donnant deux fragments du gène rapporteur.

Les deux fragments ont été insérés entre les sites Hind III et Sac I du vecteur pGEM7 modifié et vérifiés pour leur orientation. Cette procédure donne le vecteur nlz sans promoteur.

## **2. Fabrication de p352nlz**

La figure 5 illustre la fabrication de ce plasmide.

Un fragment PCR de 417 pb couvrant la région -352 à +65 du gène SM 22, dont la séquence se retrouve dans la séquence SEQ ID N° 1 qui comprend l'enchaînement des nucléotides - 2126 à + 4135, a été produit et inséré dans le site Hinc II rendu bout-franc du vecteur pBluescriptSK+. La séquence du fragment PCR a été vérifiée. Le fragment est récupéré par les sites Xba I et Xho I et inséré dans le plasmide pBLCAT 3 dans les sites correspondants, générant le recombinant p352CAT. Pour créer p352nlz, le fragment Xba I/XhoI de p352CAT a été inséré dans les sites correspondants de nlz.

## **3. Fabrication de p2126nlz**

Pour créer p2126nlz, un fragment Bsp 120I/Sph I de 1,9 kb recouvrant la région -2126 à -206 du locus SM 22, dont la séquence se retrouve dans la séquence SEQ ID N° 1 qui comprend l'enchaînement des nucléotides - 2126 à + 4135, a été inséré dans les sites respectifs de p352nlz. Ce plasmide a été déposé le 25 mars 1996 auprès de la CNCM sous le n° I-1685. Sa fabrication est illustrée par la figure 6.

#### **4. Fabrication de p2126INTnlz**

La fabrication de ce plasmide est illustrée par la figure 7.

Un fragment Bsp 120I/Hind III de 5,8 kb recouvrant la région -2126 à +3651 du locus SM 22 a été inséré dans les sites correspondants de pnlz, donnant pINT-nlz. Ensuite un fragment PCR de 496 pb de la base +3648 à +4143 du locus SM 22 a été généré en utilisant une amorce en amont et une amorce en aval conçue pour introduire un site Hind III en position 4135-4140. Ce produit de PCR a été digéré avec Hind III et lié dans le site Hind III de pINT-nlz pour donner p2126INT-nlz. La construction p2126INTnlz contient donc 2126 paires de bases de la région en 5' du début de transcription, le premier exon, le premier intron entier et 15 paires de bases du deuxième exon avec les quatre dernières paires de bases mutées de CATG en GCTT pour éliminer le codon d'initiation de traduction et introduire un site Hind III comme mentionné ci-dessus. La partie de l'insert dérivé du promoteur de SM 22 contenu dans le plasmide p2126 INTnlz est constitué par la séquence SEQ ID N°1. Ce plasmide a été déposé le 25 mars 1996 auprès de la CNCM sous le n°I-1686. Sa fabrication est illustrée par la figure 7.

#### **EXEMPLE 3 :**

#### **Expression des constructions p2126 nlz et p2126 INTnlz dans les animaux transgéniques**

Deux constructions utilisant le gène reporteur lacZ avec signal de localisation nucléaire du virus SV 40 ont été fabriquées, comme décrit dans l'exemple 2.

La première construction, p2126nlz contient 2126 bp de la région en amont et le premier exon (65 bp). La seconde construction p2126INTnlz est identique à la première à l'exception de l'addition du premier intron et des premières 12bp du second exon.

Comme indiqué dans la partie Matériel et Méthodes, chacune des constructions a donné naissance à au moins une souris transgénique exprimant le gène reporteur.

Les embryons obtenus de 12,5 à 15,5 jours après le coit (dpc) ont été colorés de manière à mettre en évidence l'activité  $\beta$ gal. Les expressions des deux constructions sont identiques, à tous les stades



observés. bien que p2126INTnlz donne une coloration un petit peu plus intense.

L'expression a été détectée en premier lieu au jour 8,5 dans la région du coeur de l'embryon et dans les vaisseaux de l'enveloppe  
5 cardiaque. Au jour 9,5, l'expression de lacZ dans le coeur augmente et est confinée au ventricule droit, aux artères.

Entre 12,5 et 15,5 jours une ségrégation du marquage en fonction des régions devient évident, le marquage étant confiné au ventricule droit (figures 8A et 8B). Les coupes montrent que l'expression a lieu  
10 dans le myocarde du ventricule droit, avec des marquages mineurs dans le ventricule gauche et dans l'adrium droit (figure 9A).

Cette expression dans le coeur est seulement transitoire, et n'apparaît pas chez l'adulte. Elle est vraisemblablement inhibée dans les phases tardives du développement des foetus. L'expression précoce est  
15 aussi observée dans les segments du rostre à partir de 9,5 jours qui s'étend aux segments caudaux à 10,5 jours, et atteint des niveaux importants à 11,5 jours.

L'expression diminue à partir de 12,5 jours dans les segments du rostre (figure 8B) et finalement n'est plus détectable à 14,5 jours (figure  
20 10E).

Les coupes montrent que l'expression du transgène est restreinte à la région myotomale des segments (figure 9 D).

La plus grande partie de l'expression a lieu dans le système vasculaire de l'embryon en développement. A partir de 9,5 jours,  
25 l'expression est détectée dans l'aorte dorsale. A 10,5 jours, la coloration de l'aorte dorsale et des coudes aortiques augmente et à 11,5 jours, l'expression lacZ est clairement visible dans l'aorte dorsale, les coudes aortiques, les artères iliaques, les artères ombilicales, les artères de la carotide et dans les vaisseaux majeurs de la tête.

30 L'expression du transgène dans les vaisseaux augmente à partir de 12,5 jours, date à laquelle les vaisseaux intercostaux peuvent être facilement vus (Figure 8A).

A 15.5 jours (figure 8B), lacZ est présent dans les vaisseaux majeurs, incluant ceux des membres et de la queue, et la coloration du  
35 tronc pulmonaire devient visible. L'expression dans le système

vasculaire persiste chez l'adulte et est clairement visible dans l'aorte, dans le tronc pulmonaire et dans l'artère pulmonaire droite (figure 10A), ainsi que dans les vaisseaux des intestins (figure 10B), dans la vessie et dans l'utérus.

5 Des coupes histologiques révèlent mieux les différences dans l'expression du transgène dans les artères et dans les veines. La section de l'embryon à 12,5 jours (figure 9A) montre un niveau d'expression très élevée dans les couches musculaires des artères de la carotide, de la quatrième artère du coude de l'aorte, du tronc pulmonaire proximal et de  
10 la partie proximale de l'aorte ascendante, tandis que les parties gauche et droite des veines cardinales antérieures restent non colorées. L'absence d'expression du transgène dans les veines musculaires ressort à l'évidence des coupes du ductus venosus, ainsi que des veines portales et ombilicales. La même situation est observée chez l'adulte, où  
15 l'expression ne peut jamais être détectée dans les veines caves (figure 10A) ni dans les veines pulmonaires.

Ces observations mettent en évidence des différences qui n'avaient jusqu'alors pas été mises en évidence entre les couches de muscles lisses des artères et des veines. D'autres vues des artères  
20 montrent que l'expression est restreinte à la couche de muscles et est absente de l'endothélium (figure 9C).

Il est de plus évident, au vu de l'ensemble des observations (figures 8A, 8B), que l'expression du transgène est absente des muscles lisses des viscères. Ceci est inattendu, car le gène SM 22 endogène est  
25 supposé être exprimé dans les tissus musculaires lisses vasculaires et des viscères. Dans les coupes histologiques de la région abdominale d'embryon à 14,5 jours, les couches de muscles de l'estomac, de l'intestin et de la vessie en développement sont facilement reconnaissables (figure 9C). L'expression du transgène dans cette  
30 région est clairement détectée dans les artères ombilicales, mais pas dans les couches de muscles des viscères de l'estomac, du duodenum, de l'intestin grêle et du gros intestin, et de la vessie, où aucune expression lacZ n'est observée. En outre, aucune expression n'est observée dans l'oesophage et dans la trachée (figure 9A), ainsi que  
35 dans les bronches du poumon.

La coloration du tissu musculaire de l'adulte donne essentiellement les mêmes résultats. La figure 10B montre que l'expression du transgène est absente dans les muscles du colon, alors que le vaisseau mésentère est très fortement coloré par lacZ. Par contre, la coloration  
5 par immunofluorescence de la coupe d'une région intestinale similaire montre que la protéine SM 22 endogène est présente dans les deux tissus (figure 10C). L'absence d'expression du transgène dans les muscles lisses des viscères de l'adulte est de plus confirmée par des colorations de l'oesophage, de la trachée, de la vessie, du canal  
10 déférent et de l'utérus pour l'activité  $\beta$ gal. En aucun cas, il ne peut être détectée une expression  $\beta$ gal dans les couches musculaires de ces tissus.

Curieusement, une coloration importante est observée dans le cytoplasme des cellules épithéliales du lumen du canal déférent et dans  
15 l'épithélium des conduits rénaux. Ceci est vraisemblablement dû à une activité galactosidase endogène, la coloration étant aussi visible dans les animaux non transgéniques témoins.

#### **EXEMPLE 4 :**

**Fabrication d'une construction comportant des régions**  
**régulatrices du gène SM 22 et le gène codant la thymidine kinase du**  
20 **virus de l'herpès (SM 22 INT-TK).**

Le fragment HIV-LTR (-167, + 80) est délété du plasmide pLTR-TK par digestion Hind III dont les extrémités sont rendues bout franc avec l'enzyme de Klenow, suivie d'une digestion Xho I. Le plasmide  
25 refermé en insérant un linker contenant les sites XhoI en 5', les sites Hind III et Bsp 120I dans l'ordre 5'→3' et une extrémité 3' bout-franc pour donner le vecteur p-TK. Un fragment Bsp 120I/Hind III de 5.8 kb recouvrant la région -2126 à + 3648 du locus SM 22 est inséré dans p-TK pour obtenir SM 22 INT-TK. Ce fragment est compris dans  
30 l'enchaînement de nucléotides de la SEQ ID N°1.

Cette construction est schématiquement représentée sur la figure  
11.

#### **EXEMPLE 5:**

**Fabrication du plasmide p445 nlz portant le fragment -445 à +**  
35 **65 du gène SM22 fusionné au gène lacZ, et obtention d transgènes**

## **Matériels et méthodes**

### **Construction de p445nlz**

Le plasmide p2126nlz a été coupé par Xba I qui a un site unique dans le « polylinker » en 5' en amont de la séquence SM 22. Les  
5 extrémités 5' protudantes du fragment ont été remplies à l'aide de la polymérase Klenow, par du dNTP.

Après précipitation le fragment a été coupé par Pst I qui a un site unique en position -445 relativement au nucléotide +1 de début de la transcription.

10 Les extrémités 3' protudantes de la coupure Pst I ont été hydrolysées à la polymérase Klenow sans dNTP. Le fragment comprenant les séquences plasmidiques plus la séquence SM 22 de -445 à +65 fusionnée au gène lacZ a été purifié à partir d'un gel d'agarose et refermé par la ligase T4.

15 Pour les expériences de transgénèse le fragment injecté dans les oeufs de souris a été purifié à partir d'une double digestion Pst I/Nsi I du plasmide p2126nlz.

Les résultats obtenus avec cette construction et ceux obtenus avec p2126 INTlacZ et p2126lacZ pour comparaison, figurent dans le tableau  
20 ci-après.

Les propriétés de régulation transcriptionnelle de la séquence du gène SM 22 entre les nucléotides -445 et +65 ont été analysées in vivo dans des souris transgéniques par deux méthodes. La première méthode appelée analyse transitoire consiste à analyser le patron  
25 d'expression du gène lacZ directement dans les embryons fondateurs FO à un stade donné. La deuxième méthode consiste à établir des lignées stables à partir de fondateurs FO adultes.

### **1. Analyse transitoire**

L'analyse transitoire a permis d'obtenir quatre embryons ayant  
30 intégré le transgène p445nlz dans leur génome sur un total de 13 embryons. Parmi les quatre embryons positifs analysés à 12,5 jours de développement, un embryon n'exprimait pas le transgène. Deux embryons présentaient un patron d'expression identique à celui de la lignée p2126 INT nlz au même stade, c'est-à-dire spécifique des cellules  
35 musculaires lisses artérielles, et enfin un troisième embryon présentait

les mêmes territoires d'expression artérielle additionnés de divers sites d'expression ectopique essentiellement diffus dans le mésenchyme embryonnaire.

## **2. Lignées stables**

5 Deux lignées stables indépendantes ont été établies à partir d'un mâle (lignée n°1) et d'une femelle (lignée n°9).

L'expression du transgène a été analysée dans l'embryon au stade 15.5 jours pour la lignée 1 et 17,5 jours de développement dans la lignée 9. Les deux lignées présentent le même type d'expression dans  
10 les artères: au niveau du coeur, le transgène est exprimé dans l'aorte dorsale et le tronc pulmonaire artériel.

A la différence des lignées obtenues avec les transgènes p2126nlz et p2126 INTnlz, les deux lignées p445nlz expriment lacZ très fortement dans les parois de l'oesophage et de la trachée respiratoire  
15 qui sont composées de muscle lisse à ce stade.

Les deux lignées présentent des différences entre elles dans certaines régions de l'embryon. Pour la lignée n°1, le transgène s'exprime fortement dans les bronchioles en continuité de la trachée à 15,5 jours ce qui n'est pas observé dans la lignée n°9 à 17,5 jours où  
20 lacZ n'est trouvé que dans la trachée. Cette différence peut être due soit à une modification de la régulation due à des sites d'intégration différents dans les deux lignées, soit à la différence de stade qui fait que l'expression de lacZ peut être réprimée entre le jour 15,5 et 17,5.

Dans les deux lignées on trouve donc une expression du transgène  
25 dans des organes qui contiennent une forte composante de cellules musculaires lisses. Il est intéressant de noter qu'il s'agit d'une partie au moins des cellules musculaires lisses de type respiratoire, la trachée, et une partie des cellules musculaires lisses de type viscéral, l'oesophage.

Enfin des sites d'expression ectopiques différents, c'est-à-dire où la  
30 présence de la protéine ou de l'ARN SM 22 endogène n'a pas été reportée, ont été trouvés dans les deux lignées. La lignée 1 présente entre autre une expression forte très localisée dans les deux doigts postérieurs de chaque membre. La lignée 9 exprime fortement lacZ dans les régions antérieures du cerveau.

L'expression de cette construction à deux stades de l'évolution est illustrée par les figures 12A à 12D qui représentent des embryons de 15,5 jours p.c. (12A, 12B, 12C) et un adulte de 2 mois (12 D) SM 445 nlz colorés pour l'activité  $\beta$ -galactosidase. On note l'expression dans le tissu musculaire lisse des artères, de l'oesophage et des bronches sur les figures 12A, B, C ainsi que dans des sites ectopiques comme la moëlle épinière (12B) ou les deux doigts postérieurs de chaque membre (12A). On note aussi l'absence d'expression dans l'aorte thoracique ou le tronc pulmonaire artériel chez l'adulte (12D). Seule une expression dans la trachée est maintenue.

Les résultats obtenus montrent qu'une séquence de 450 paires de bases est suffisante pour garder la même expression dans l'embryon et le fœtus mais pas chez l'adulte.

TABLEAU

<i>Constructions</i>	<i>N° de naissances</i>	<i>Nomenclature</i>	<i>Expression Embryon</i>	<i>Adulte</i>
SM 22 p 2126INTlacZ	6			
	22	M9 F23	artères.	aorte, artère pulmonaire
SM27 p2126lacZ	17	M3 M4	artères ventricule droit	aorte, artère pulmonaire ventricule droit
		F9		
SM445 p445lacZ	8	n°1	artère+no sp	
		n°6	artères seules	
	5	n°7		
	17	n°1	artères seules	
		M1	artère+no sp	pas d'expression
		F9	artère + no sp	pas d'expression

## LISTE DE SEQUENCES

## (1) INFORMATIONS GENERALES:

## (i) DEPOSANT:

(A) NOM: INSTITUT PASTEUR  
(B) RUE: 28 RUE DU DOCTEUR ROUX  
(C) VILLE: PARIS  
(E) PAYS: FRANCE  
(F) CODE POSTAL: 75015

(A) NOM: UNIVERSITE PARIS 7  
(B) RUE: 1 PLACE JUSSIEU  
(C) VILLE: PARIS  
(E) PAYS: FRANCE  
(F) CODE POSTAL: 75005

(ii) TITRE DE L' INVENTION: SEQUENCES EN AMONT DU GENE SM22

(iii) NOMBRE DE SEQUENCES: 5

## (iv) FORME DECHIFFRABLE PAR ORDINATEUR:

(A) TYPE DE SUPPORT: Floppy disk  
(B) ORDINATEUR: IBM PC compatible  
(C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS  
(D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

## INFORMATIONS POUR LA SEQ ID NO: 1:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 6261 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: double  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

## (vi) ORIGINE:

(A) ORGANISME: SOURIS



GG GCCCCAGGAA  
TGIGTTTCCT TCTCTCCACC ATGTTTTTAT AGCTCTTGGG  
CTGGGAGAAG AGGCGGGTCT  
GGGTCTTTGT TTCTGAGCTT TGTCTATGT TCCTCCATGC  
TACGGTTGCA ATTGTTTTCT  
ATGAACGAGT ACATTCAATA AAGACAACCA GACCTGGGAT  
TTGGGGTCTT ACTGATGTGT  
TGGGAGGTGC AGGAGCCTCC GTGTCCCATT TATTTTGGCC  
TTCCCGTCTC GTTTCTGTGC  
GTGGCTACAT TGGGAATGAC CTTCTTGAT CCCACCAAGC  
CACCCATTGA TTCTGTAAAC  
ATGTGACCCT TGCTCCAAGC ATTGCTTACA GGAGCAGGAT  
ACTGAAAGTG TGTCTGTGCC  
CTCTCCTGAT ACCCCCTCCC TTCAGCAGGC ACACAGCACC  
TGACTACCCA CCACGTATGT  
AAACGTCAGT ATCCTTTCCA GCCAGCTCTG CAGATGGGTG  
TCCAGGCTGT GCATGATGCA  
CCTCAAGTGG GCAGAGCTTG CAGGCCAAGG TTTTAAAGGC  
TGTTCAAGAA TGGATGGCAA  
GCAGGATCTA AGAGGAGGGG GGGTGTGTGT TGTTTGGGGG  
GGGGGTGGTT TTGGTTTGT  
TTTTTTGAGA CAGGGTTTCT CTGIGTGGCC CTGGCCCTCC  
TGGAACCCAC TCTGTAGACC  
AGGCTGGCCT TGAATCAGA AATCTGCCTG CCTCTGCCTC  
CCGAGTGCTG GGATTAAAGG  
CGTGTGCCCA TCGAGGAGGG AGATTTTATT TAGATTATAA  
AAAGGACGGG ATTTGGGGAA  
TCCTGTCTAG TGAATTCAGG ACGTAATCAG TGGCTGGGAA  
GCAAGAGCTC TAGAGGAGCT  
CCAGCTTATT ATGACCCTTC CTTAGATGC CACAAGGAGG  
TGCTGGAGTT CTATGCACCA  
ATAGCTTAAA CCAGCCAGGC TGGCTGTAGT GGATTGAGCG  
TCTGAGGCTG CACCTCTCTG  
GCCTGCAGCC AGTTCTGGGT GAGACTGACC CTGCCTGAGG  
GTTCTCTCCT TCCCTCTCTC  
TACTCCTTCC TCCCTCTCCC TCTCCCTCTC TCTGTTTCCT GAGGTTTCCA  
GAATTGGGGA  
TGGGACTCAG AGACACCACT AAAGCCTTAC CTTTTAAGAA  
GTTGCATTCA GTGAGTGTGT  
GAGACATAGC ACAGATAGGG GCAGAGGAGA GCTGGTTCTG  
TCTCCACTGT GTTTGGTCTT  
GGGTACTGAA CTCAGACCAT CAGGTGTGAT AGCAGTTGTC  
TTTAACCCTA ACCCTGAGCC  
TGTCTCACCT GTCCCTTCCC AAGACCACTG AAGCTAGGTG  
CAAGATAAGT GGGGACCCTT  
TCTGAGGTGG TAGGATCTTT CACGATAAGG ACTATTTTGA  
AGGGAGGGAG GGTGACACTG  
TCCTAGTCCT CTTACCCTAG TGTCTCCAGC CTTGCCAGGC  
CTTAAACATC CGCCCATTGT  
CACCGCTCTA GAAGGGGCCA CCCTTGACTT GCTGCTAAAC  
AAGGCACTCC CTAGAGAAGA

GTCCACTGTA GGCAGATAGG TGACAGGTGG CAGATAGGTG  
ACAGATAGGT GACAGGTGGA  
GGAGCTTTGG AACTGGGACT GGACAGCCCT GGGACCCTGT  
TCCTCCCAA GGGTCTTGGT  
GGTTCCTT GGGGCTCTCT AAAGGATGTC AGTGGGCTGT  
TGCCACATCT ATATAAGAGG  
ACTAGTCTTC TGGAATTTAG GTGTGATCTC TCAGGGATGC  
AGAAATGCTC ACCCTTACTG  
TCATTTTATG GGCTGAGGTA CCACAGGCAG ATATACCCTG  
GTCTGCTTGT TGTCCAGGGT  
CTCTGCTACA TGGAGGCCCT TTTCCACAGC CTAACCTCTC  
TACCTGCTGA CAGGAGGGCT  
GGATGGCCAC AGGCATCCAA CGTGCGCATC ATGCAGGTGT  
TTTGCGTTGG AGCTTTTGT  
TAGAAATACC CTGGTGGGCT GCCAAACCAC CACCCATATC  
CCTCTCTCT CTCTGCTGCC  
TCTAAGATGA CAGCTTGATT TTTCTTATAG TGATTTTTTT TTTTGGTTTT  
GTTTTTTTGT  
TTGTTTTAAG TTAGCATACA AAGTAATACA TTTCATCATG  
GCATTGGAC ATACATATAT  
ATTTTATTTG CTCTCCTGGC CTCTTCTCAA AGAGACTTCT CTGGACTTTC  
TTGTATTTTT  
GGTTGTGAGC CTAGCCTTTA ACGGCTGAGC CATCTCTCCA  
GCCCTTCTTT GGACTTTCTA  
CTTCATACTT CCCACCAGTC TGGGAAGAAG GGCACATGGA  
ATCTTGAGAG CATGACCTGA  
CCCAGACCTG ACAGATGTCA AGGCTGCAGT GTATGCTCTT  
GTTTCGTACGG CTTGTTCTTA  
GTCCTGCAGT TCAGAACTTT CTGGAGACTG AGAAGTGCAT  
GTGAGGACAC TCTCCTCCCA  
TCTTTTCCTC TAGTGGCTAG TGATGTTTGG TTTTTTGTTT  
TGAGACAGGG TTTCTCTGTA  
TAGCCCTAGC TATCCTGGAA CTCACCTTGT AGATCAGGCT  
GGCCTCCAAC TCAGAAATCT  
GCCTGCCTCT GCCTCCCGAG TGCTGGGACT AAAGGCGTGC  
GCCACCACTG TCCAGTCAGG  
AGTAGAAGGA AACTGTAAGG TGCTTGAGAC AGGCTGAGTA  
GAGGCTAGGA GGAAGGGGCA  
CCGCAGTCAC CGGCTCCATG ACTCTGTGAC TTTTGTGGTT  
CCTTGTCGCA GCGGTTCTG  
GTGGTGGTGG TGGTCGGGGG TTGGGGGGAG GGGGCAGGCC  
ACACAGTGGG GTGTGGGAGG  
GAATAGCTGT TGACAACTTC CCAACAGAAA CCAGGCTTTT  
GAGTCCTCA GGGTAGCTTG  
AGAGGGTACT CAGAAAGCCG TGTCCATGTC CCCTTTCCTT  
CACCTCAGGG AAGTAAGTTG  
CCTATAGGGT TGTCAATTCA ATGAGGTCTT CTGGTTATTC TGTTTTTCTC  
TCAATGTTGG  
TGTTGGGCTC AGGGAATGCT TTGGAGAAGG TGGTGGGAAC  
TGGAGAAGGG AAGATCAGTT

TACCATACCT GTGGGCAGGA TGACCCATGT TCTGCCATGC  
ACTTGGTAGC CTTGGAAAGG  
CCACTTTGAA CCTCAATTTT CTCAACTGTT AAATGGAGTG  
GTAAC TGCTA TCTCATAATA  
AAGGGGAACG TGAGGAAGGC GTTTGGATAG TGCCTGGTTG  
CGGCCAGGCT GCAGTCAAGA  
CTAGTTCCCA CCAACTCGAT TTAAAGCCT TGCAAGAAGG  
TGGCTTGTTT GTCCCTTGCA  
GGTTCCTTTG CTCGGGCCAA ACTCTAGAAT GCCTCCCCCT  
TTCTTTCTCA TTGAAGAGCA  
GACCCAAGTC CGGGTAACAA GGAAGGGTTT CAGGGTCCTG  
CCCATAAAAG GTTTTTCCCG  
GCCGCCCTCA GCACCGCCCC GCCCCGACCC CCGCAGCATC  
TCCAAAGCAT GCAGAGAATG  
TCICCGGCTG CCCCCGACAG ACTGCTCCAA CTTGGTGTCT  
TTCCCCAAAT ATGGAGCCTG  
TGTTGAGTGA GTGGGGCGGC CCGGGGTGGT GAGCCAAGCA  
GACTTCCATG GGCAGGGAGG  
GGCGCCACGG GCGGCAGAG GGGTGACATC ACTGCCTAGG  
CGGCCTTAA ACCCCTCACC  
CAGCCGGCGC CCCGGCCCGT CTGCCCCAGC CCAGACACCG  
AAGCTACTCT CCTTCCAGTC  
CACAAACGAC CAAGCCTTGT AAGTGCAAGT CATGGGAGCA  
GAAGGGCTGT GGGCTCAATT  
AGATCCCCTA GTCTCTTCTA GTTTGCTGGG TGGAATTGGG  
TCCCTAGAGA CCATTCTCTG  
TGTTAGACAA AAAGTCTGGG TTAAAATGCC TAGGATGATT  
TGA CTGGGGC AAAAGAATAA  
ATGGGGTGAG AGGGAGGCTC AAATTCAGTC ACTGTCCCAC  
CCATAGGTGT ATGGGCTATG  
TGTTAGGCC AAAAGAGGTGA CAAATGAGGC CAAGGGAACA  
ACTCCATCTT TGGATCTCCA  
AGAAGGTGAG GGGCTAAGTT CTGGAAAGCA GTGACCCACT  
GATGGTCCCC AGGGCTAATG  
CAACTCGGGG GAGCCAGGAG GTAGCCCCCT CAGGCAGTGG  
AGGACTAAAG ATCTTATTTT  
TTGTAGCGCT AGGGATCAAA CCCCAGGGCG CTATGTGTGG  
CAGGCATGTG CTCCATCTAC  
CACAGAAGTT TAATCCTTCA GACTAGCCTG GGATAGGGCC  
TGCTTTTTCT TTCCTTTTCT  
CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTTT  
CTTTCTTTT  
CTCTCTTTCA CTCTCTCTTT CTAATTTCTT TTTCTTTTTT TCTTTCTTTT  
CTTTAGACAG  
GGTTTCTCTG TGTAGCCCTG GCTGTTCTGG AACTCACTCT  
TTAGACCAGG CTGGCCTCGA  
ATCTCAGAAA TCTACCTGCC TCTGCCTCCC AAGTGCTGGG  
ATTAAAGGCG TGTGCCACCA  
CTGCCAGCT AAGGTTTGCT TTTTGATGGC AGCTTGGTCC  
AGTTTGAAAG TAGGAGGTCA

TACTGTGTAA ACTCACTGGT TAAAGTACCC CCTCCCTTCC  
ACCCTGCAAT ACACACACAC  
ACACACACAC ACACACACAC ACACACACAC ACCCCATCTC  
GAAGAGCTCT ATTAAGCTCC  
AGGTGCACTG TAGTTCACAG ACTGCATCTT CCAGGTTTGC  
TCCCACTTCA CAAGCAGAGA  
ACTCATAACT GAAGGGGGTG ACAGCACAGG GGAAGGGAAA  
GCAAGATGTT TAGAGTCTGA  
CAGCTGGCCC GGGACCAGAG CCATGTGGTA ATGTTTGCTC  
CACTCCCATC CACCTCCACG  
GCTGTGATGT GGAGAAGGTC CCCGCTTTCA TGGGAAGGAG  
GTGGGGGAGC CTGTCACTG  
CTCCATGCTC ACACAATTTT TCTCTCAACC AATGACCTCT  
CAGAAGCAGG GGTGGGTTG  
CAAAATTCTT CAGATACCTC AACAGATGGC ATCCCACTCA  
GGCTATCCCT GCTGACTAGG  
TCTGGCTCCA GCCCTGACTG TATCTACCCA GGGACCTACC  
TGCCTGCTTT GCTCCTATAG  
CCTTCCTCCG TGTCTGGGTC CCCAGAGAGC TGCCGGCATA  
GGCCTTTGAG GCAACAGCTG  
GCATACAGGC CAGGCTTCCC ATGCTCTGGC TAGCAGATTC  
TCTGCCCTGG AGGACTTTGA  
CTGCATGGTT TCTCTCACTG CTGCAACAGT CAGAGCTGGC  
CCACACGGGC ACAACAGCGC  
ACTTCCATCT GGGTCTCCCT GAGAATGCCG CTGTTTTCTG  
AGAACCCTTG GACTCTGGTG  
GCTTTATCAG GTCTTTTTGT CAGCTGCGCT TTGGGGGATG  
AACTTTGCTC TTCTGGCTTC  
TGGGTCAGAG GGTAAGATT TGGTGGCAAC CGGTAGCTAG  
AGAAAGATAG CTA CTGGCTG  
AATTTGGAGG ACATGGCTTC TGGAAAACCT CTCTAGTGCT  
TTTCTGGCTA GTCTTGGCAA  
AGTAAAAATG CTCTGATAGC CAGCCCGGGT GATGCAGGGC  
TTCCTGTTCT AGGCCTTTCT  
GTACAAAATT AGTGAGACAT TGCCTCAAAA CTATGAAACA  
AGCCAGACTC TGTTGAAGCA  
CGCCTTTAAT CCCAGCACTC AGGAGGCTGA AGCAGGCAAG  
ATCTCTGTTA GTTGGAGGCC  
AGTCTACAGG AAAGTTCTAC AACAGCAGAG GCCAGACAGT  
GCAACCCTTT CTGGGGGTGT  
GGGGGAGGAA AACCCAACAA AAACACAAAC TATAAAACAA  
AGAGAAGGCC GAGGACAAAG  
CTTAGCAATG CATACTTCCC TTTCTATGTG AAGCCCTGGG  
CTCCACCAGT ACTGCAGAAA  
GAAGCAAGCA ATGAGGGACA GGAGGTTGGC TCTAGGCCCA  
GGGGTTGTCA AAATAGTCCA  
CAGGCCAAAG GCAGCCTGAT GTCTGTTTTT ATAAACAAAA  
TTTTATTGGC ACACATTGGT  
TATGTATCAG CTAGGCTATT TTCATTACAA TAGAGGCCAT  
ATGGTCTGTA AAGTCTAAAA

TATTTACTCT GCTGTTTTAC ATAAAAAGTT GACAGACTCT  
TGCTCTAGAC TGACAAATAT  
CTAAGACCTT GTTTTCTGAG GTTCAAGTTT CAGAGGGGTC  
TCTGCAGCAA GTGGGTAAAG  
CTGGTCTAGG TCATGCTATG ATGTCTAGGG TCCCCTCAGA  
GTGGAAGGCC TGCTTAGCAC  
AAATGAAGTA AAGTAACTTG CTGGCTCTTT GTTCTTTTCT  
CCACACTCTA TACTTTAGCT  
CTGCCTC

3 =

## (2) INFORMATIONS POUR LA SEQ ID NO: 2:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 2190 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(vi) ORIGINE:

(A) ORGANISME: SOURIS

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:2 :

GGGCCCCAGG AATGTGTTTC CTTCTCTCCA CCATGTTTTT ATAGCTCTTG GGCTGGGAGA  
AGAGGCGGGT CTGGGTCTTT GTTCTGAGC TTTGTTCTAT GTTCCTCCAT GCTACGGTTG 1  
CAATTGTTTT CTATGAACGA GTACATTCAA TAAAGACAAC CAGACCTGGG ATTTGGGGTC 1

TTACTGATGT GTTGGGAGGT GCAGGAGCCT CCGTGTCCCA TTTATTTTGG CCTTCCCGTC	2.
TCGTTTCTGT GCGTGGCTAC ATTGGGAATG ACCTTCCTTG ATCCCACCAA GCCACCCATT	30
GATTCTGTAA ACATGTGACC CTTGCTCCAA GCATTGCTTA CAGGAGCAGG ATACTGAAAG	38
TGTGTCTGTG CCCTCTCCTG ATAACCCCTC CCTTCAGCAG GCACACAGCA CCTGACTACC	46
CACCACGTAT GTAAACGTCA GTATCCTTTC CAGCCAGCTC TGCAGATGGG TGTCCAGGCT	48
GTGCATGATG CACCTCAAGT GGGCAGAGCT TGCAGGCCAA GGTTTTAAAG GCTGTTTCAGG	54
AATGGATGGC AAGCAGGATC TAAGAGGAGG GGGGGTTGTT GTTGTTTGGG GGGGGGGTGG	60
TTTTGGTTTG TTTTTTTTGA GACAGGGTTT CTCTGTGTGG CCCTGGCCCT CCTGGAACCC	66
ACTCTGTAGA CCAGGCTGGC CTTGAACTCA GAAATCTGCC TGCCTCTGCC TCCCGAGTGC	72
TGGGATTAAA GCGTGTGCC CATCGAGGAG GGAGATTTTA TTTAGATTAT AAAAAGGACG	78
GGATTTGGGG AATCCTGTCT AGTGAATTCA GGACGTAATC AGTGGCTGGG AAGCAAGAGC	84
TCTAGAGGAG CTCCAGCTTA TTATGACCCT TCCTTCAGAT GCCACAAGGA GGTGCTGGAG	90
TTCTATGCAC CAATAGCTTA AACCAGCCAG GCTGGCTGTA GTGGATTGAG CGTCTGAGGC	96
TGCACCTCTC TGGCCTGCAG CCAGTTCTGG GTGAGACTGA CCCTGCCTGA GGGTCTCTCTC	102
CTTCCCTCTC TCTACTCCTT CCTCCCTCTC CCTCTCCCTC TCTCTGTTTC CTGAGGTTTC	108
CAGAATTGGG GATGGGACTC AGAGACACCA CTAAAGCCTT ACCTTTTAAG AAGTTGCATT	114
CAGTGAGTGT GTGAGACATA GCACAGATAG GGGCAGAGGA GAGCTGGTTC TGTCTCCACT	120
GTGTTTGGTC TTGGGTACTG AACTCAGACC ATCAGGTGTG ATAGCAGTTG TCTTTAACCC	126
TAACCCTGAG CCTGTCTCAC CTGTCCCTTC CCAAGACCAC TGAAGCTAGG TGCAAGATAA	132
GTGGGGACCC TTTCTGAGGT GGTAGGATCT TTCACGATAA GGACTATTTT GAAGGGAGGG	138
AGGGTGACAC TGTCTAGTC CTCTTACCCT AGTGTCTCCA GCCTTGCCAG GCCTTAAACA	144
TCCGCCCATT GTCACCGCTC TAGAAGGGGC CACCCTTGAC TTGCTGCTAA ACAAGGCACT	150
CCCTAGAGAA GATACCATAC CTGTGGGCAG GATGACCCAT GTTCTGCCAT GCACTTGGTA	156
GCCTTGGAAG GGCCACTTTG AACCTCAATT TTCTCAACTG TTAAATGGAG TGGTAACTGC	162
TATCTCATAA TAAAGGGGAA CGTGAGGAAG GCGTTTGGAT AGTGCCTGGT TCGGGCCAGG	168
CTGCAGTCAA GACTAGTTCC CACCAACTCG ATTTTAAAGC CTTGCAAGAA GGTGGCTTGT	174

TTGTCCCTTG CAGGTTTCCTT TGCTCGGGCC AACTCTAGA ATGCCTCCCC CTTTCTTTCT	180
CATTGAAGAG CAGACCCAAG TCCGGGTAAC AAGGAAGGGT TTCAGGGTCC TGCCCATAAA	185
AGGTTTTTCC CGGCCGCCCT CAGCACCGCC CCGCCCCGAC CCCC GCAGCA TCTCCAAAGC	190
ATGCAGAGAA TGTCTCCGGC TGCCCCCGAC AGACTGCTCC AACTTGGTGT CTTTCCCCAA	195
ATATGGAGCC TGTGTGGAGT GAGTGGGGCG GCCCGGGTG GTGAGCCAAG CAGACTTCCA	200
TGGGCAGGGA GGGGCGCCAC GGGGCGGCAG AGGGGTGACA TCACTGCCTA GCGGCCTTT	205
AAACCCCTCA CCCAGCCGGC GCCCCGGCCC GTCTGCCCCA GCCCAGACAC CGAAGCTACT	210
CTCCTTCCAG TCCACAAACG ACCAAGCCTT	215



## (2) INFORMATIONS POUR LA SEQ ID NO: 3

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1661 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(vi) ORIGINE:

(A) ORGANISME: SOURIS

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3

GGGCCCCAGG AATGTGTTTC CTTCTCTCCA CCATGTTTTT ATAGCTCTTG GGCTGGGAGA  
AGAGGCGGGT CTGGGTCTTT GTTCTGAGC TTTGTTCTAT GTTCCTCCAT GCTACGGTTG  
CAATTGTTTT CTATGAACGA GTACATTCAA TAAAGACAAC CAGACCTGGG ATTTGGGGTC

TTACTGATGT GTTGGGAGGT GCAGGAGCCT CCGTGTCCCA TTTATTTTGG CCTTCCCGTC	2
TCGTTTCTGT GCGTGGCTAC ATTGGGAATG ACCTTCCTTG ATCCCACCAA GCCACCCATT	3
GATTCTGTAA ACATGTGACC CTTGCTCCAA GCATTGCTTA CAGGAGCAGG ATACTGAAAG	3
TGTGTCTGTG CCCTCTCCTG ATAACCCCTC CCTTCAGCAG GCACACAGCA CCTGACTACC	4
CACCACGTAT GTAAACGTCA GTATCCTTTC CAGCCAGCTC TGCAGATGGG TGTCCAGGCT	4
GTGCATGATG CACCTCAAGT GGGCAGAGCT TGCAGGCCAA GGTTTTAAAG GCTGTTCAGG	5
AATGGATGGC AAGCAGGATC TAAGAGGAGG GGGGGTTGTT GTTGTTTGGG GGGGGGGTGG	6
TTTTGGTTTG TTTTTTTTGA GACAGGGTTT CTCTGTGTGG CCCTGGCCCT CCTGGAACCC	6
ACTCTGTAGA CCAGGCTGGC CTTGAACTCA GAAATCTGCC TGCCTCTGCC TCCCGAGTGC	7
TGGGATTAAA GGCCTGTGCC CATCGAGGAG GGAGATTTTA TTTAGATTAT AAAAAGGACG	7
GGATTTGGGG AATCCTGTCT AGTGAATTCA GGACGTAATC AGTGGCTGGG AAGCAAGAGC	8
TCTAGAGGAG CTCCAGCTTA TTATGACCCT TCCTTCAGAT GCCACAAGGA GGTGCTGGAG	9
TTCTATGCAC CAATAGCTTA AACCAGCCAG GCTGGCTGTA GTGGATTGAG CGTCTGAGGC	9
TGCACCTCTC TGGCCTGCAG CCAGTTCTGG GTGAGACTGA CCCTGCCTGA GGGTTCTCTC	10
CTTCCCTCTC TCTACTCCTT CCTCCCTCTC CCTCTCCCTC TCTCTGTTTC CTGAGGTTTC	10
CAGAATTGGG GATGGGACTC AGAGACACCA CTAAAGCCTT ACCTTTTAAG AAGTTGCATT	11
CAGTGAGTGT GTGAGACATA GCACAGATAG GGGCAGAGGA GAGCTGGTTC TGTCTCCACT	12
GTGTTTGGTC TTGGGTACTG AACTCAGACC ATCAGGTGTG ATAGCAGTTG TCTTTAACCC	12
TAACCCTGAG CCTGTCTCAC CTGTCCCTTC CCAAGACCAC TGAAGCTAGG TGCAAGATAA	13
GTGGGGACCC TTTCTGAGGT GGTAGGATCT TTCACGATAA GGACTATTTT GAAGGGAGGG	13
AGGGTGACAC TGTCCTAGTC CTCTTACCCT AGTGTCTCCA GCCTTGCCAG GCCTTAAACA	14
TCCGCCCATT GTCACCGCTC TAGAAGGGGC CACCCTTGAC TTGCTGCTAA ACAAGGCACT	15
CCCTAGAGAA GATACCATAC CTGTGGGCAG GATGACCCAT GTTCTGCCAT GCACTTG GTA	15
GCCTTGAAAA GGCCACTTTG AACCTCAATT TTCTCAACTG TTAAATGGAG TGGTAACTGC	16
TATCTCATAA TAAAGGGGAA CGTGAGGAAG GCGTTTGGAT AGTGCCTGGT TGCGGCCAGG	16

## INFORMATIONS POUR LA SEQ ID NO: 4

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: . paires de bases
- (B) TYPE : nucléotide
- (C) NOMBRE DE BRINS : double
- (D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADN (génomique)

(iii) HYPOTHETIQUE : NON

(vi) ORIGINE :

(A) ORGANISME : SOURIS

TCTAGAGGCC ACGGAACGGT GCCAAGCACA CAGTCCCTTT TGCCTCTTC  
ACGGGAGCAG  
GAGTCCCACT GCCTGTCGTG GAAAGGGAGG AACATGCCAG GTCCCTGTGT  
GTCCTTGGCC  
CTGTCTCACC AAAGGACTCA GGGCTGGTTT CTGAGTTTCC GTCCAGTATT  
TAGCCAAGTC  
CTGTGTTAGT CACGTAGGCC TAAGAGCCTT GGCGTTTACA GAGTCACCCA  
GCTCTGGCCC  
CTGGCATTCT GGTCTTGGC GTTTACAGAG TCACCCAGCT CCAGGCCCT  
GGCACTTTGG  
TACTTGGTTG CCCTTCACTC CACCAGGTCC ATTCCAGATG CCAAGAGTGG  
GCCCCAGGAA  
TGTGTTTCT TCTCTCCACC ATGTTTTAT AGCTCTTGGG CTGGGAGAAG  
AGGCGGGTCT  
GGGTCTTTGT TTCTGAGCTT TGTCTATGT TCCTCCATGC TACGGTTGCA  
ATTGTTTTCT  
ATGAACGAGT ACATTCAATA AAGACAACCA GACCTGGGAT TTGGGGTCTT  
ACTGATGTGT  
TGGGAGGTGC AGGAGCCTCC GTGTCCATT TATTTTGGCC TTCCCGTCTC  
GTTTCTGTGC  
GTGGCTACAT TGGGAATGAC CTTCTTGAT CCCACCAAGC CACCCATTGA  
TTCTGTAAAC  
ATGTGACCCT TGCTCCAAGC ATTGCTTACA GGAGCAGGAT ACTGAAAGTG  
TGTCTGTGCC  
CTCTCCTGAT AACCCCTCCC TTCAGCAGGC ACACAGCACC TGACTIONCA  
CCACGTATGT  
AAACGTCACT ATCCTTTCCA GCCAGCTCTG CAGATGGGTG TCCAGGCTGT  
GCATGATGCA  
CCTCAAGTGG GCAGAGCTTG CAGGCCAAGG TTTTAAAGGC TGTTCAAGAA  
TGGATGGCAA  
GCAGGATCTA AGAGGAGGGG GGGTTGTTGT TGTTGGGGG GGGGGTGGTT  
TTGGTTTGT  
TTTTTTGAGA CAGGGTTTCT CTGTGTGGCC CTGGCCCTCC TGAACCCAC  
TCTGTAGACC  
AGGCTGGCCT TGAATCAGA AATCTGCCTG CCTCTGCCTC CCGAGTGCTG  
GGATTAAAGG  
CGTGTGCCCC TCGAGGAGGG AGATTTTATT TAGATTATAA AAAGGACGGG  
ATTGGGGAA  
TCCTGTCTAG TGAATTCAGG ACGTAATCAG TGGCTGGGAA GCAAGAGCTC  
TAGAGGAGCT  
CCAGCTTATT ATGACCCTTC CTTGAGATGC CACAAGGAGG TGCTGGAGTT  
CTATGCACCA  
ATAGCTTAAA CCAGCCAGGC TGGCTGTAGT GGATTGAGCG TCTGAGGCTG  
CACCTCTCTG  
GCCTGCAGCC AGTTCTGGGT GAGACTGACC CTGCCTGAGG GTTCTCTCCT  
TCCCTCTCTC  
TACTCCTTCC TCCCTCTCCC TCTCCCTCTC TCTGTTTCTT GAGGTTTCCA  
GAATTGGGGA  
TGGGACTCAG AGACACCACT AAAGCCTTAC CTTTTAAGAA GTTGCATTCA  
GTGAGTGTGT  
GAGACATAGC ACAGATAGGG GCAGAGGAGA GCTGGTTCTG TCTCCACTGT  
GTTTGGTCTT  
GGGTACTGAA CTCAGACCAT CAGGTGTGAT AGCAGTTGTC TTTAACCCTA  
ACCCTGAGCC  
TGTCTCACCT GTCCCTTCCC AAGACCACTG AAGCTAGGTG CAAGATAAGT  
GGGGACCCTT  
TCTGAGGTGG TAGGATCTTT CACGATAAGG ACTATTTTGA AGGGAGGGAG  
GGTGACACTG

TCCTAGTCCT CTTACCCTAG TGTCTCCAGC CTTGCCAGGC CTAAACATC  
CGCCCATTGT  
CACCGCTCTA GAAGGGGCCA CCCTTGACTT GCTGCTAAAC AAGGCACTCC  
CTAGAGAAGA  
TACCATACCT GTGGGCAGGA TGACCCATGT TCTGCCATGC ACTTGGTAGC  
CTTGGAAAGG  
CCACTTTGAA CCTCAATTT CTCAACTGTT AAATGGAGTG GTAAGTGCTA  
TCTCATAATA  
AAGGGGAACG TGAGGAAGGC GTTTGGATAG TGCCTGGTTG CGGCCAGGCT  
GCAGTCAAGA  
CTAGTTCCCA CCAACTCGAT TTAAAGCCT TGCAAGAAGG TGGCTTGT  
GTCCCTTGCA  
GGTTCCTTTG CTCGGGCCAA ACTCTAGAAT GCCTCCCCCT TTCTTTCTCA  
TTGAAGAGCA  
GACCCAAGTC CGGGTAACAA GGAAGGGTTT CAGGGTCCTG CCCATAAAG  
GTTTTTCCCG  
GCCGCCCTCA GCACCGCCCC GCCCGACCC CCGCAGCATC TCCAAAGCAT  
GCAGAGAATG  
TCTCCGGCTG CCCCCGACAG ACTGCTCAA CTTGGTGTCT TTCCCCAAAT  
ATGGAGCCTG  
TGTGGAGTGA GTGGGECGGC CCGGGTGGT GAGCCAAGCA GACTTCCATG  
GGCAGGGAGG  
GGCGCCACGG GCGGCAGAG GGGTGACATC ACTGCCTAGG CGGCCTTAA  
ACCCCTCACC  
CAGCCGGCGC CCCGGCCCGT CTGCCCCAGC CCAGACACCG AAGTACTCT  
CCTTCCAGTC  
CACAAACGAC CAAGCCTTGT AAGTGCAAGT CATGGGAGCA GAAGGGCTGT  
GGGCTCAATT  
AGATCCCCTA GTCTCTTCTA GTTTGCTGGG TGGAATTGGG TCCCTAGAGA  
CCATTCTCTG  
TGTTAGACAA AAAGTCTGGG TTAAATGCC TAGGATGATT TGAAGGGG  
AAAAGAATAA  
ATGGGGTGAG AGGGAGGCTC AAATTCAGTC ACTGTCCAC CCATAGGTGT  
ATGGGGCTATG  
TGTTAGGCCC AAAGAGGTGA CAAATGAGGC CAAGGGAACA ACTCCATCTT  
TGGATCTCCA  
AGAAGGTGAG GGGCTAAGTT CTGGAAGCA GTGACCCACT GATGGTCCCC  
AGGGCTAATG  
CAACTCGGGG GAGCCAGGAG GTAGCCCCCT CAGGCAGTGG AGGACTAAAG  
ATCTTATTTT  
TTGTAGCGCT AGGGATCAA CCCAGGGCG CTATGTGTGG CAGGCATGTG  
CTCCATCTAC  
CACAGAAGTT TAATCCTTCA GACTAGCCTG GGATAGGGCC TGCTTTTCT  
TTCCTTTTCT  
CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTTT  
CTTTCCTTTT  
CTCTCTTTCA CTCTCTCTTT CTAATTTCTT TTTCTTTT TCTTTCTTT CTTAGACAG  
GGTTTCTCTG TGAGCCCTG GCTGTTCTGG AACTCACTCT TTAGACCAGG  
CTGGCCTCGA  
ATCTCAGAAA TCTACCTGCC TCTGCCTCCC AAGTGCTGGG ATTAAGGGCG  
TGTGCCACCA  
CTGCCAGCT AAGGTTTGCT TTTTGATGGC AGCTTGGTCC AGTTGAAAG  
TAGGAGGTCA  
GTCCACTGTA GGCAGATAGG TGACAGGTGG CAGATAGGTG ACAGATAGGT  
GACAGGTGGA  
GGAGCTTTGG AACTGGGACT GGACAGCCCT GGGACCCTGT TCCTCCCAA  
GGGTCTTGGT

GGTTCCCTT GGGGCTCTT AAAGGATGTC AGTGGGCTGT TGCCACATCT  
ATATAAGAGG  
ACTAGTCTTC TGGAATTTAG GTGTGATCTC TCAGGGATGC AGAAATGCTC  
ACCCTTACTG  
TCATTTTATG GGCTGAGGTA CCACAGGCAG ATATACCCTG GTCTGCTGT  
TGTCAGGGT  
CTCTGCTACA TGGAGGCCCT TTTCCACAGC CTAACCTCTC TACCTGCTGA  
CAGGAGGGCT  
GGATGGCCAC AGGCATCCAA CGTGCGCATC ATGCAGGTGT TTTGCGTTGG  
AGCTTTTGT  
TAGAAATACC CTGGTGGGCT GCCAAACCAC CACCCATATC CCTCTCTCT  
CTCTGCTGCC  
TCTAAGATGA CAGCTTGATT TTTCTTATAG TGATTTTTT TTTTGGTTT GTTTTTTGT  
TTGTTTTAAG TTAGCATACA AAGTAATACA TTTCATCATG GCATTTGGAC  
ATACATATAT  
ATTTTATTTG CTCTCCTGGC CTCTTCTCAA AGAGACTTCT CTGGACTTTC  
TTGTATTTT  
GGTTGTGAGC CTAGCCTTTA ACGGCTGAGC CATCTCTCCA GCCCTTCTTT  
GGACTTTCTA  
CTTCATACTT CCCACCAGTC TGGGAAGAAG GGCACATGGA ATCTTGAGAG  
CATGACCTGA  
CCCAGACCTG ACAGATGTCA AGGCTGCAGT GTATGCTCTT GTTCGTACGG  
CTTGTTCTTA  
GTCCTGCAGT TCAGAACTTT CTGGAGACTG AGAAGTGCAT GTGAGGACAC  
TCTCTCCCA  
TCTTTCTCTC TAGTGGCTAG TGATGTTTGG TTTTTGTTT TGAGACAGGG  
TTTCTCTGTA  
TAGCCCTAGC TATCCTGGAA CTCACTTTGT AGATCAGGCT GGCCTCCAAC  
TCAGAAATCT  
GCCTGCCCTCT GCCTCCCGAG TGCTGGGACT AAAGGCGTGC GCCACCACTG  
TCCAGTCAGG  
AGTAGAAGGA AACTGTAAGG TGCTTGAGAC AGGCTGAGTA GAGGCTAGGA  
GGAAGGGGCA  
CCGCAGTCAC CGGCTCCATG ACTCTGTGAC TTTTGTGGTT CCTTGTGCGA  
GCGGTTCTG  
GTGGTGGTGG TGGTCGGGGG TTGGGGGGAG GGGGCAGGCC ACACAGTGGG  
GTGTGGGAGG  
GAATAGCTGT TGACAACTTC CCAACAGAAA CCAGGCTTTT GAGTCCTCCA  
GGGTAGCTTG  
AGAGGGTACT CAGAAAGCCG TGTCCATGTC CCCTTTCCTT CACCTCAGG  
AAGTAAGTTG  
CCTATAGGGT TGTCATTTCA ATGAGGTCTT CTGGTTATTC TGTTTTTCTC  
TCAATGTTGG  
TGTTGGGCTC AGGGAATGCT TTGGAGAAGG TGGTGGAAC TGGAGAAGGG  
AAGATCAGTT  
TACTGTGTAA ACTCACTGGT TAAAGTACCC CCTCCCTTCC ACCCTGCAAT  
ACACACACAC  
ACACACACAC ACACACACAC ACACACACAC ACCCATCTC GAAGAGCTCT  
ATTAAGCTCC  
AGGTGCACTG TAGTTCACAG ACTGCATCTT CCAGGTTTGC TCCCACTTCA  
CAAGCAGAGA  
ACTCATAACT GAAGEGGGTG ACAGCACAGG GGAAGGGAAA GCAAGATGTT  
TAGAGTCTGA  
CAGCTGGCCC GGGACCAGAG CCATGTGGTA ATGTTTGCTC CACTCCCATC  
CACCTCCACG  
GCTGTGATGT GGAGAAGGTC CCCGCTTTC TGGGAAGGAG GTGGGGGAGC  
CTGTCATCTG

CTCCATGCTC ACACAATTTT TCTCTCAACC AATGACCTCT CAGAAGCAGG  
GGTTGGTTTG  
CAAAATTCTT CAGATACCTC AACAGATGGC ATCCCACTCA GGCTATCCCT  
GCTGACTAGG  
TCTGGCTCCA GCCCTGACTG TATCTACCCA GGGACCTACC TGCCTGCTTT  
GCTCCTATAG  
CCTTCCTCCG TGTCTGGGTC CCCAGAGAGC TGCCGGCATA GGCTTTGAG  
GCAACAGCTG  
GCATACAGGC CAGGCTTCCC ATGCTCTGGC TAGCAGATTG TCTGCCCTGG  
AGGACTTTGA  
CTGCATGGTT TCTCTCACTG CTGCAACAGT CAGAGCTGGC CCACACGGGC  
ACAACAGCGC  
ACTTCCATCT GGGTCTCCCT GAGAATGCCG CTGTTTTCTG AGAACCCTG  
GACTCTGGTG  
GCTTTATCAG GTCTTTTGT CAGCTGCGCT TTGGGGGATG AACTTTGCTC  
TTCTGGCTTC  
TGGGTCAGAG GGTAAAGATT TGGTGGCAAC CGGTAGCTAG AGAAAGATAG  
CTACTGGCTG  
AATTTGGAGG ACATGGCTTC TGGAAAACCT CTCTAGTGCT TTTCTGGCTA  
GTCTTGGCAA  
AGTAAAAATG CTCTGATAGC CAGCCCGGGT GATGCAGGGC TTCCTGTTGG  
AGGCCTTTCT  
GTACAAAATT AGTGAGACAT TGCCTCAAAA CTATGAAACA AGCCAGACTC  
TGTGGAAGCA  
CGCCTTTAAT CCCAGCACTC AGGAGGCTGA AGCAGGCAAG ATCTCTGTTA  
GTTGGAGGCC  
AGTCTACAGG AAAGTTCTAC AACAGCAGAG GCCAGACAGT GCAACCCTTT  
CTGGGGGTGT  
GGGGGAGGAA AACCCAACAA AAACACAAAC TATAAAACAA AGAGAAGGCC  
GAGGACAAAG  
CTTAGCAATG CATACTTCCC TTTCTATGTG AAGCCCTGGG CTCCACCAGT  
ACTGCAGAAA  
GAAGCAAGCA ATGAGGGACA GGAGGTTGGC TCTAGGCCCA GGGGTTGTCA  
AAATAGTCCA  
CAGGCCAAAG GCAGCCTGAT GTCTGTTTTT ATAAACAAAA TTTTATTGGC  
ACACATTGGT  
TATGTATCAG CTAGGCTATT TTCATTACAA TAGAGGCCAT ATGGTCTGTA  
AAGTCTAAAA  
TATTTACTCT GCTGTTTTAC ATAAAAAGTT GACAGACTCT TGCTCTAGAC  
TGACAAATAT  
CTAAGACCTT GTTTTCTGAG GTTCAAGTTT CAGAGGGGTC TCTGCAGCAA  
GTGGGTAAAG  
CTGGTCTAGG TCATGCTATG ATGTCTAGGG TCCCCTCAGA GTGGAAGGCC  
TGCTTAGCAC  
AAATGAAGTA AAGTAACTTG CTGGCTCTTT GTTCTTTTCT CCACACTCTA  
TACTTTAGCT  
CTGCCTCAAC ATG

45

## (2) INFORMATIONS POUR LA SEQ ID NO: 5

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 785 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: double
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(vi) ORIGINE:

(A) ORGANISME: SOURIS

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO:5

```

GGGCCCCAGG AATGTGTTTC CTTCTCTCCA CCATGTTTTT ATAGCTCTTG GGCTGGGAGA      8
AGAGGCGGGT CTGGGTCTTT GTTCTGAGC TTTGTTCTAT GTTCCTCCAT GCTACGGTTG      12
CAATTGTTTT CTATGAACGA GTACATTCAA TAAAGACAAC CAGACCTGGG ATTTGGGGTC      16
TTACTGATGT GTTGGGAGGT GCAGGAGCCT CCGTGTCCTTA TTTATTTTGG CCTTCCCGTC      24
TCGTTTCTGT GCGTGGCTAC ATTGGAATG ACCTTCCTTG ATCCCACCAA GCCACCCATT      30
GATTCTGTAA ACATGTGACC CTTGCTCCAA GCATTGCTTA CAGGAGCAGG ATACTGAAAG      36
TGTGTCTGTG CCCTCTCCTG ATAACCCCTC CCTTCAGCAG GCACACAGCA CCTGACTACC      42
CACCACGTAT GTAAACGTCA GTATCCTTTC CAGCCAGCTC TGCAGATGGG TGTCCAGGCT      48
GTGCATGATG CACCTCAAGT GGGCAGAGCT TGCAGGCCAA GGTTTTAAAG GCTGTTCAGG      54
AATGGATGGC AAGCAGGATC TAAGAGGAGG GGGGGTTGTT GTTGTGTTGGG GGGGGGGTGG      60
TTTTGGTTTG TTTTTTTTGA GACAGGGTTT CTCTGTGTGG CCCTGGCCCT CCTGGAACCC      66
ACTCTGTAGA CCAGGCTGGC CTTGAACTCA GAAATCTGCC TGCCTCTGCC TCCCGAGTGC      72
TGGGATTAAA GGCCTGTGCC CATCGAGGAG GGAGATTTTA TTTAGATTAT AAAAAGGACG      78
GGATT,                                     78

```



## References

- Almendral, J. M., Santaren, J. F., Perera, J., Zerial, M. and Bravo, R. (1989). Expression, cloning and cDNA sequence of a fibroblast serum-regulated gene encoding a putative actin-associated protein (p27). *Exp. Cell Res.*, 181, 518-530.
- Babij, P., Kelly, C. and Periasamy, M. (1991). Characterisation of a mammalian smooth muscle myosin heavy-chain gene: Complete nucleotide and protein coding sequence and analysis of the 5' end of the gene. *Proc. Natl. Acad. Sci.*, 88, 10676-10680.
- Birnstiel, M. L., Busslinger, M. and Strub, K. (1985). Transcription termination and 3' processing: the end is in site. *Cell*, 41, 349-359.
- Bour, B. A., O'Brien, M. A., Lockwood, W. L., Goldstein, E. S., Bodmer, R., Taghert, P. H., Abmayr S. M. and Nguyen, H. T. (1995). *Drosophila* MEF2, a transcription factor that is essential for myogenesis. *Genes & Dev.*, 9, 730-741.
- Chamley-Campbell, J. H. and Campbell, G. R. (1981). *Artherosclerosis*, 40, 347-357.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, 18, 5294-5299.

- Duband, J.-L., Gimona, M., Scatena, M., Sartore, S. and Small J. V. (1993). Calponin and SM 22 as differentiation markers of smooth muscle: spatiotemporal distribution during avian embryonic development. *Differentiation*, 55, 1-11.
- Dynan, W. S. and Tijan, R. (1983) The promoter-specific transcription factor SP 1 binds to upstream sequences in the SV 40 early promoter. *Cell*, 35, 79-87.
- Frohman, M. A., Dush, M. K. and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci.*, 85, 8998-9002.
- Gimona, M., Sparrow, M. P., Strasser, P., Herzog, M. and Small, J. V. (1992). Calponin and SM 22 isoforms in avian and mammalian smooth muscle. Absence of phosphorylation in vivo. *Eur. J. Biochem.* 205, 1067-1075.
- Kelly, R., Alonso, S., Tajbakhsh, S., Cossu, G. and Buckingham, M. (1995). Myosin light chain 3F regulatory sequences confer regionalized cardiac and skeletal muscle expression in transgenic mice. *J. of Cell Biol.*, 129/2, 383-396.
- Gorman, C.M., Moffat L.F. and B.H.H. (1982). Recombinant genomes which express chloramphenicol acetyl-transferase in mammalian cells. *Mol. Cell Biol.* 2: 1044-1051.
- Kemp, P. R., Osbourn, J. K., Grainger D. J. and Metcalfe, J. C. (1995). Cloning and analysis of the promoter region of the rat SM 22 $\alpha$  gene. *Biochem. J.*, 310, 1037-1043.

- Kim, J.-H., Bush I, P. R. and Kumar, C. C. (1993). Smooth muscle  $\alpha$ -actin promoter activity is induced by serum stimulation of fibroblast cells. *Biochem. Biophys. Res. Com.*, 190/3, 1115-1121.
- Kramerov, D. A., Lekakh, I. V., Samarina, O. P. and Ryskov, A. P. (1982). The sequences homologous to major interspersed repeats B1 and B2 of mouse genome are present in mRNA and small cytoplasmic poly(A)<sup>+</sup> RNA. *Nuc. Acid Res.*, 10/23, 7477-7491.
- Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov A. P., Bayev, A. A. and Georgiev, G. P. (1980). The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA. *Nuc. Acid Res.*, 8/6, 1201-1215.
- Lees-Miller, J. P., Heeley, D. H., Smillie, L. B. and Kay, C. M. (1987a). Isolation and characterization of an abundant and novel 22-kDa protein (SM 22) from chicken gizzard smooth muscle. *J. Biol. Chem.*, 262/7, 2988-2993.
- Lees-Miller, J. P., Heeley, D. H. and Smillie, L. B. (1987b). An abundant and novel protein of 22 kDa (SM 22) is widely distributed in smooth muscles. *Biochem. J.*, 244, 705-709.
- Li, Z., Marchand, P., Humbert, J., Babinet C. and Paulin D. (1993). Desmin sequence elements regulating skeletal muscle-specific expression in transgenic mice. *Development*, 117/3, 947-959.

Lilly, B., Zha, B., Rangasayakulu, G., Paterson, B. M., Schulz, R. A. and Olson, E. N. (1995). Requirement for MADS Domain transcription factor D-MEF2 for muscle formation in *Drosophila*. *Science*, 267, 688-693.

Lückow, B. and Schütz, G. (1987). CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. *Nuc. Acid Res.*, 15/13, 5490.

Min, B., Foster, D. N. and Strauch, A. R. (1990). The 5'-flanking region of the mouse vascular smooth muscle  $\alpha$ -actin gene contains evolutionarily conserved sequence motifs within a functional promoter. *J. Biol. Chem.*, 265/27, 16667-16675.

Miwa, T., Manabe, Y., Kurokawa, K., Kamada, S., Kanda, N., Bruns, G., Ueyama, H. and Kakunaga, T. (1991). Structure, chromosome location and expression of the human smooth muscle (enteric type)  $\gamma$ -actin gene: Evolution of six human actin genes. *Mol. Cell. Biol.*, 11/6, 3296-3306.

Mössler H. (1995). Characterization of the gene structure and the promoter region of the smooth muscle specific protein SM 22 of *mus musculus*. PhD thesis at the University of Salzburg, Austria.

Nishida, W., Kitami, Y., Abe, M. and Hiwada, K. (1991). Gene cloning and nucleotide sequence of SM 22 $\alpha$  from the chicken gizzard smooth muscle. *Biochem. Intern.*, 23/4, 663-668.

- Olson, E. N. (1990). My D family: a paradigm for development?. *Genes and Dev.*, 4, 1454-1461.
- Osbourne, J. K., Weissberg, P. L. and Shanahan, C. M. (1995). A regulatory element downstream of the rat SM 22a gene transcription start point enhances reporter gene expression in vascular smooth muscle cells. *Gene*, 154, 249-253.
- Pearlstone, J. R., Weber, M., Lees-Miller, J. P., Carpenter, M. R. and Smillie, L. B. (1987). Amino acid sequence of chicken gizzard smooth muscle SM 22 $\alpha$ . *J. Biol. Chem.*, 262/13, 5985-5991.
- Prinjha, R. K., Shapland, C. E., Hsuan, J. J., Totty, N. F., Mason, I. J. and Lawson, D. (1994). Cloning and sequencing of cDNAs encoding the actin cross-linking protein transgelin defines a new family of actin-associated proteins. *Cell Motil. and Cytoskel.*, 28, 243-255.
- Reddy, S., Ozgur, K., Lu, M., Chang, W., Mohan, S., Kumar, C. and Ruley, H. E. (1990). Structure of the human smooth muscle  $\alpha$ -actin gene. *J. Biol. Chem.*, 265/3, 1683-1687.
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci.*, 74, 5463ff.
- Santaren, J. F., Blüthmann, H., MacDonald-Bravo, H. and Bravo, R. (1987). Specific antibody against a protein (p27) present in nonestablished fibroblasts. A putative Microfilament associated protein. *Exp. Cell. Res.*, 173, 341-348.

- Shanahan, C. M., Weissberg, P. L. and Metcalf, J. C. (1993). Isolation of gene markers of differentiated and proliferating vascular smooth muscle cells. *Circ. Res.*, 73, 193 - 204.
- Shanahan, C., Cary, N. R. B., Metcalf, J. C. and Weissberg, P. L. (1994). High Expression of genes for calcification-regulating proteins in human atherosclerotic plaques. *J. Clin. Invest.*, 93, 2393 - 2402.
- Shapland, C., Lowings, P. and Lawson, D. (1988). Identification of new actin associated polypeptides that are modified by viral transformation and changes in cell shape. *J. Cell Biol.*, 107, 153-161.
- Shapland, C., Hsuan, J. J. Totty, N. F. and Lawson, D. (1993). Purification and properties of Transgelin: A transformation and shape change sensitive actin-gelling protein. *J. Cell Biol.*, 121/5, 1065-1073.
- Solway, J., Seltzer, J., Samaha, F. F., Kim, S., Alger, L. E., Niu, Q., Morrissey, E. E., Ip, H. S. and Parmacek, M. S. (1995). Structure and expression of a smooth muscle cell-specific gene, SM 22 $\alpha$ . *J. of Biol. Chem.*, 270/22, 13460-13469.
- Thweatt, R., Lumpkin, C. K. and Goldstein, S. (1992). A novel gene encoding a smooth muscle protein is overexpressed in senescent human fibroblasts. *Biochem. Biophys. Res. Com.*, 187/1, 1-7.

Tzeng, Y.-J., Guhl, E., Graessmann, M. and Graessmann, A. (1993). Breast cancer formation in transgenic animals induced by the whey acidic protein SV 40 T antigen (WAP-SV-T) hybrid gene. *Oncogene*, 8, 1965-1971.

LI LI, J.M. MIANO, B. MERCER and E. OLSON

J. OF CELL BIOLOGY 1996 - VOL 132 N° 5, p. 849-859

## REVENDICATIONS

1. Séquence d'ADN caractérisée en ce qu'elle comprend :

- un fragment de la séquence en amont de la partie codante du  
5 gène de la protéine SM22, ou d'une séquence hybridant dans des  
conditions de forte stringence à ladite séquence en amont, ledit  
fragment étant susceptible d'induire une expression spécifique d'un  
gène dans des cellules eucaryotes, et
- une séquence codant pour une protéine, ou un ARN, d'intérêt  
10 thérapeutique.

2. Séquence d'ADN selon la revendication 1, caractérisée en ce  
qu'elle comprend :

- un fragment de la séquence en amont de la partie codante du  
gène de la protéine SM 22, ou d'une séquence hybridant dans des  
15 conditions de forte stringence à ladite séquence en amont, ledit  
fragment étant susceptible d'induire une expression spécifique in vivo  
d'un gène dans les cellules des artères, et
- une séquence codant pour une protéine, ou un ARN, d'intérêt  
thérapeutique.

20 3. Séquence selon l'une des revendications 1 ou 2, caractérisée en  
ce que ladite protéine , ou ledit ARN, est susceptible d'inhiber la  
croissance des cellules du muscle lisse, d'activer la croissance des  
cellules endothéliales, de consolider les parois des artères et/ou  
d'induire une réponse immunitaire.

25 4. Séquence selon l'une des revendications 1 à 3, caractérisée en  
ce que ledit fragment est compris dans la séquence située entre les  
nucléotides -2126 et +4135 du gène de souris de la protéine SM22  
(SEQ ID n° 1.)

30 5. Séquence selon l'une des revendications 1 à 4, caractérisée en  
ce que ledit fragment est compris dans la séquence située entre les  
nucléotides - 2126 à +65 du gène de souris de la protéine SM 22 (SEQ  
ID N°2).

6. Séquence selon l'une des revendications 1 à 5 caractérisée en  
ce qu'elle contient au moins une partie de la séquence située entre les



nucléotides -2126 et -445 du gène de la souris de la protéine SM 22 (SEQ ID N° 3).

7. Séquence selon l'une des revendications 1 à 6, caractérisée en ce que ladite protéine d'intérêt thérapeutique est une protéine induisant la formation d'un composé cytotoxique.

8. Séquence selon la revendication 7, caractérisée en ce que ladite protéine est la thymidine kinase du virus de l'herpès.

9. Séquence selon l'une des revendications 1 à 6, caractérisée en ce que ladite protéine d'intérêt thérapeutique présente un effet cytostatique.

10. Séquence selon l'une des revendications 1 à 6 caractérisée en ce que ladite protéine d'intérêt thérapeutique présente une activité lipolytique.

11. Séquence selon la revendication 10, caractérisée en ce que ladite protéine est la lipoprotéine lipase.

12. Séquence selon l'une des revendications 1 à 6, caractérisée en ce que ladite protéine d'intérêt thérapeutique est un facteur de croissance des cellules endothéliales.

13. Séquence selon la revendication 11, caractérisée en ce que ledit facteur est une interleukine.

14. Séquence selon l'une des revendications 1 à 6, caractérisée en ce que ladite protéine d'intérêt thérapeutique est une protéine musculaire ou de structure tissulaire.

15. Séquence selon l'une des revendications 1 à 6, caractérisée en ce que l'ARN d'intérêt thérapeutique est l'ARN antisens de la protéine p53.

16. Séquence d'ADN située en amont de la partie codante du gène de souris de la protéine SM 22 caractérisée en ce qu'elle comprend au moins une partie de la séquence SEQ ID N°5 suivante, ou une séquence s'hybridant avec celle-ci dans des conditions stringentes :

GGGCCCCAGG	AATGTGTTTC	CTTCTCTCCA	CCATGTTTTT
ATAGCTCTTG	GGCTGGGAGA	AGAGGCGGGT	CTGGGTCTTT
GTTTCTGAGC	TTTGTTCAT	GTTCTCCAT	GCTACGGTTG
CAATTGTTTT	CTATGAACGA	GTACATTCAA	TAAAGACAAC

	CAGACCTGGG	ATTTGGGGTC	TTACTGATGT	GTTGGGAGGT
	GCAGGAGCCT	CCGTGTCCCA	TTTATTTTGG	CCTTCCCGTC
	TCGTTTCTGT	GCGTGGCTAC	ATTGGGAATG	ACCTTCCTTG
	ATCCCACCAA	GCCACCCATT	GATTCTGTAA	ACATGTGACC
5	CTTGCTCCAA	GCATTGCTTA	CAGGAGCAGG	ATACTGAAAG
	TGTGTCTGTG	CCCTCTCCTG	ATAACCCCTC	CCTTCAGCAG
	GCACACAGCA	CCTGACTACC	CACCACGTAT	GTAAACGTCA
	GTATCCTTTC	CAGCCAGCTC	TGCAGATGGG	TGTCCAGGCT
	GTGCATGATG	CACCTCAAGT	GGGCAGAGCT	TGCAGGCCAA
10	GGTTTTAAAG	GCTGTTCAGG	AATGGATGGC	AAGCAGGATC
	TAAGAGGAGG	GGGGGTTGTT	GTTGTTTGGG	GGGGGGGTGG
	TTTTGGTTTG	TTTTTTTTGA	GACAGGGTTT	CTCTGTGTGG
	CCCTGGCCCT	CCTGGAACCC	ACTCTGTAGA	CCAGGCTGGC
	CTTGAACTCA	GAAATCTGCC	TGCCTCTGCC	TCCCGAGTGC
15	TGGGATTAAA	GGCGTGTGCC	CATCGAGGAG	GGAGATTTTA
	TTTAGATTAT	AAAAAGGACG	GGATT	

17. Séquence d'ADN située en amont de la partie codante du gène  
de la souris de la protéine SM22, caractérisée en ce qu'elle comprend  
20 au moins une partie de la séquence SEQ ID n° 1 suivante, ou une  
séquence s'hybridant avec celle-ci dans des conditions stringentes :

GG GCCCCAGGAA  
TGTGTTTCCT TCTCTCCACC ATGTTTTTAT AGCTCTTGGG  
CTGGGAGAAG AGGCGGGTCT  
GGGTCTTTGT TTCTGAGCTT TGTTCTATGT TCCTCCATGC  
TACGGTTGCA ATTGTTTTCT  
ATGAACGAGT ACATTCAATA AAGACAACCA GACCTGGGAT  
TTGGGGTCTT ACTGATGTGT  
TGGGAGGTGC AGGAGCCTCC GTGTCCCATT TATTTTGGCC  
TTCCCGTCTC GTTTCTGTGC  
GTGGCTACAT TGGGAATGAC CTTCTTGAT CCCACCAAGC  
CACCCATTGA TTCTGTAAAC  
ATGTGACCCT TGCTCCAAGC ATTGCTTACA GGAGCAGGAT  
ACTGAAAGTG TGTCTGTGCC  
CTCTCCTGAT AACCCCTCCC TTCAGCAGGC ACACAGCACC  
TGACTACCCA CCACGTATGT  
AAACGTCAGT ATCCTTTCCA GCCAGCTCTG CAGATGGGTG  
TCCAGGCTGT GCATGATGCA  
CCTCAAGTGG GCAGAGCTTG CAGGCCAAGG TTTTAAAGGC  
TGTTCAAGAA TGGATGGCAA  
GCAGGATCTA AGAGGAGGGG GGGTTGTTGT TGTTTGGGGG  
GGGGGTGGTT TTGGTTTGT  
TTTTTTGAGA CAGGGTTTCT CTGTGTGGCC CTGGCCCTCC  
TGGAACCCAC TCTGTAGACC  
AGGCTGGCCT TGAATCAGA AATCTGCCTG CCTCTGCCTC  
CCGAGTGCTG GGATTAAAGG  
CGTGTGCCCA TCGAGGAGGG AGATTTTATT TAGATTATAA  
AAAGGACGGG ATTTGGGGAA  
TCCTGTCTAG TGAATTCAGG ACGTAATCAG TGGCTGGGAA  
GCAAGAGCTC TAGAGGAGCT  
CCAGCTTATT ATGACCCTTC CTTCAGATGC CACAAGGAGG  
TGCTGGAGTT CTATGCACCA  
ATAGCTTAAA CCAGCCAGGC TGGCTGTAGT GGATTGAGCG  
TCTGAGGCTG CACCTCTCTG  
GCCTGCAGCC AGTTCTGGGT GAGACTGACC CTGCCTGAGG  
GTTCTCTCCT TCCCTCTCTC  
TACTCCTTCC TCCCTCTCCC TCTCCCTCTC TCTGTTTCCT GAGGTTTCCA  
GAATTGGGGA  
TGGGACTCAG AGACACCACT AAAGCCTTAC CTTTTAAGAA  
GTTGCATTCA GTGAGTGTGT  
GAGACATAGC ACAGATAGGG GCAGAGGAGA GCTGGTTCTG  
TCTCCACTGT GTTTGGTCTT  
GGGTACTGAA CTCAGACCAT CAGGTGTGAT AGCAGTTGTC  
TTTAACCCTA ACCCTGAGCC  
TGTCTCACCT GTCCCTTCCC AAGACCACTG AAGCTAGGTG  
CAAGATAAGT GGGGACCTT  
TCTGAGGTGG TAGGATCTTT CACGATAAGG ACTATTTTGA  
AGGGAGGGAG GGTGACACTG  
TCCTAGTCCT CTTACCCTAG TGTCTCCAGC CTTGCCAGGC  
CTTAAACATC CGCCATTGT  
CACCGCTCTA GAAGGGGCCA CCCTTGACTT GCTGCTAAAC  
AAGGCACTCC CTAGAGAAGA

GTCCACTGTA GGCAGATAGG TGACAGGTGG CAGATAGGTG  
ACAGATAGGT GACAGGTGGA  
GGAGCTTTGG AACTGGGACT GGACAGCCCT GGGACCCTGT  
TCCTCCCAA GGGTCTTGGT  
GGTTCCCTT GGGGCTCTCT AAAGGATGTC AGTGGGCTGT  
TGCCACATCT ATATAAGAGG  
ACTAGTCTTC TGAATTTAG GTGTGATCTC TCAGGGATGC  
AGAAATGCTC ACCCTTACTG  
TCATTTTATG GGCTGAGGTA CCACAGGCAG ATATACCCTG  
GTCTGCTTGT TGTCCAGGGT  
CTCTGCTACA TGGAGGCCCC TTTCCACAGC CTAACCTCTC  
TACCTGCTGA CAGGAGGGCT  
GGATGGCCAC AGGCATCCAA CGTGCGCATC ATGCAGGTGT  
TTTGCCTTGG AGCTTTTGT  
TAGAAATACC CTGGTGGGCT GCCAAACCAC CACCCATATC  
CCTCTCTCCT CTCTGCTGCC  
TCTAAGATGA CAGCTTGATT TTTCTTATAG TGATTTTTTT TTTTGGTTTT  
GTTTTTTGT  
TTGTTTTAAG TTAGCATACA AAGTAATACA TTTCATCATG  
GCATTTGGAC ATACATATAT  
ATTTTATTTG CTCTCCTGGC CTCTTCTCAA AGAGACTTCT CTGGACTTTC  
TTGTATTTTT  
GGTTGTGAGC CTAGCCTTTA ACGGCTGAGC CATCTCTCCA  
GCCCTTCTTT GGACTTTCTA  
CTTCATACTT CCCACCAGTC TGGGAAGAAG GGCACATGGA  
ATCTTGAGAG CATGACCTGA  
CCCAGACCTG ACAGATGTCA AGGCTGCAGT GTATGCTCTT  
GTTTCGTACGG CTTGTTCTTA  
GTCCTGCAGT TCAGAACTTT CTGGAGACTG AGAAGTGCAT  
GTGAGGACAC TCTCCTCCCA  
TCTTTTCCTC TAGTGGCTAG TGATGTTTGG TTTTTTGTTT  
TGAGACAGGG TTTCTCTGTA  
TAGCCCTAGC TATCCTGGAA CTCACCTTGT AGATCAGGCT  
GGCCTCCAAC TCAGAAATCT  
GCCTGCCTCT GCCTCCCGAG TGCTGGGACT AAAGGCGTGC  
GCCACCACTG TCCAGTCAGG  
AGTAGAAGGA AACTGTAAGG TGCTTGAGAC AGGCTGAGTA  
GAGGCTAGGA GGAAGGGGCA  
CCGCAGTCAC CGGCTCCATG ACTCTGTGAC TTTTGTGGTT  
CCTTGTCGCA GCGGTTCTG  
GTGGTGGTGG TGGTCGGGGG TTGGGGGGAG GGGGCAGGCC  
ACACAGTGGG GTGTGGGAGG  
GAATAGCTGT TGACAACTTC CCAACAGAAA CCAGGCTTTT  
GAGTCCTCCA GGGTAGCTTG  
AGAGGGTACT CAGAAAGCCG TGTCCATGTC CCCTTTCCTT  
CACCTCAGGG AAGTAAGTTG  
CCTATAGGGT TGTCATTTCA ATGAGGTCTT CTGGTTATTC TGTTTTTCTC  
TCAATGTTGG  
TGTTGGGCTC AGGGAATGCT TTGGAGAAGG TGGTGGGAAC  
TGGAGAAGGG AAGATCAGTT

TACCATACCT GTGGGCAGGA TGACCCATGT TCTGCCATGC  
ACTTGGTAGC CTTGGAAAGG  
CCACTTTGAA CCTCAATTTT CTCAACTGTT AAATGGAGTG  
GTAAGTCTA TCTCATAATA  
AAGGGGAACG TGAGGAAGGC GTTTGGATAG TGCCTGGTTG  
CGGCCAGGCT GCAGTCAAGA  
CTAGTTCCCA CCAACTCGAT TTTAAAGCCT TGCAAGAAGG  
TGGCTTGTTT GTCCCTTGCA  
GGTTCCTTTG CTCGGGCCAA ACTCTAGAAT GCCTCCCCCT  
TTCTTTCTCA TTGAAGAGCA  
GACCCAAGTC CGGGTAACAA GGAAGGGTTT CAGGGTCCTG  
CCCATAAAAG GTTTTTCCCG  
GCCGCCCTCA GCACCGCCCC GCCCGACCC CCGCAGCATC  
TCCAAAGCAT GCAGAGAATG  
TCTCCGGCTG CCCCCGACAG ACTGCTCCAA CTTGGTGTCT  
TTCCCCAAAT ATGGAGCCTG  
TGTGGAGTGA GTGGGGCGGC CCGGGGTGGT GAGCCAAGCA  
GACTTCCATG GGCAGGGAGG  
GGCGCCACGG GGCAGGAGAG GGGTGACATC ACTGCCTAGG  
CGGCCTTTAA ACCCCTCACC  
CAGCCGGCGC CCCGGCCCGT CTGCCCCAGC CCAGACACCG  
AAGCTACTCT CCTTCCAGTC  
CACAAACGAC CAAGCCTTGT AAGTGCAAGT CATGGGAGCA  
GAAGGGCTGT GGGCTCAATT  
AGATCCCCTA GTCTCTTCTA GTTTGCTGGG TGGAATTGGG  
TCCCTAGAGA CCATTCTCTG  
TGTTAGACAA AAAGTCTGGG TTAAAATGCC TAGGATGATT  
TGACTGGGGC AAAAGAATAA  
ATGGGGTGAG AGGGAGGCTC AAATTCAGTC ACTGTCCCAC  
CCATAGGTGT ATGGGCTATG  
TGTTAGGCCC AAAGAGGTGA CAAATGAGGC CAAGGGAACA  
ACTCCATCTT TGGATCTCCA  
AGAAGGTGAG GGGCTAAGTT CTGGAAAGCA GTGACCCACT  
GATGGTCCCC AGGGCTAATG  
CAACTCGGGG GAGCCAGGAG GTAGCCCCCT CAGGCAGTGG  
AGGACTAAAG ATCTTATTTT  
TTGTAGCGCT AGGGATCAAA CCCCAGGGCG CTATGTGTGG  
CAGGCATGTG CTCCATCTAC  
CACAGAAGTT TAATCCTTCA GACTAGCCTG GGATAGGGCC  
TGCTTTTTCT TTCCTTTTCT  
CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTTT  
CTTTCCTTTT  
CTCTCTTICA CTCTCTCTTT CTAATTTCTT TTTCTTTTTT TCTTTCTTTT  
CTTTAGACAG  
GGTTTCTCTG TGTAGCCCTG GCTGTTCTGG AACTCACTCT  
TTAGACCAGG CTGGCCTCGA  
ATCTCAGAAA TCTACCTGCC TCTGCCTCCC AAGTGCTGGG  
ATTAAAGGCG TGTGCCACCA  
CTGCCAGCT AAGGTTTGCT TTTTGATGGC AGCTTGGTCC  
AGTTTGAAAG TAGGAGGTCA

TACTGTGTAA ACTCACTGGT TAAAGTACCC CCTCCCTTCC  
ACCCTGCAAT ACACACACAC  
ACACACACAC ACACACACAC ACACACACAC ACCCCATCTC  
GAAGAGCTCT ATTAAGCTCC  
AGGTGCACTG TAGTTCACAG ACTGCATCTT CCAGGTTTGC  
TCCCACCTCA CAAGCAGAGA  
ACTCATAACT GAAGGGGGTG ACAGCACAGG GGAAGGGAAA  
GCAAGATGTT TAGAGTCTGA  
CAGCTGGCCC GGGACCAGAG CCATGTGGTA ATGTTTGCTC  
CACTCCCATC CACCTCCACG  
GCTGTGATGT GGAGAAGGTC CCCGCTTTCA TGGGAAGGAG  
GTGGGGGAGC CTGTCATCTG  
CTCCATGCTC ACACAATTTT TCTCTCAACC AATGACCTCT  
CAGAAGCAGG GGTTGGTTTG  
CAAAATTCTT CAGATACCTC AACAGATGGC ATCCCACTCA  
GGCTATCCCT GCTGACTAGG  
TCTGGCTCCA GCCCTGACTG TATCTACCCA GGGACCTACC  
TGCCTGCTTT GCTCCTATAG  
CCTTCCTCCG TGTCTGGGTC CCCAGAGAGC TGCCGGCATA  
GGCCTTTGAG GCAACAGCTG  
GCATACAGGC CAGGCTTCCC ATGCTCTGGC TAGCAGATTC  
TCTGCCCTGG AGGACTTTGA  
CTGCATGGTT TCTCTCACTG CTGCAACAGT CAGAGCTGGC  
CCACACGGGC ACAACAGCGC  
ACTTCCATCT GGGTCTCCCT GAGAATGCCG CTGTTTTCTG  
AGAACCCTTG GACTCTGGTG  
GCTTTATCAG GTCTTTTTGT CAGCTGCGCT TTGGGGGATG  
AACTTTGCTC TTCTGGCTTC  
TGGGTCAGAG GGTAAGATT TGGTGGCAAC CGGTAGCTAG  
AGAAAGATAG CTA CTGGCTG  
AATTTGGAGG ACATGGCTTC TGGAAAACCT CTCTAGTGCT  
TTTCTGGCTA GTCTTGCAA  
AGTAAAAATG CTCTGATAGC CAGCCCGGGT GATGCAGGGC  
TTCTGTTCG AGGCCTTTCT  
GTACAAAATT AGTGAGACAT TGCCTCAAAA CTATGAAACA  
AGCCAGACTC TGTGAAGCA  
CGCCTTTAAT CCCAGCACTC AGGAGGCTGA AGCAGGCAAG  
ATCTCTGTTA GTTGGAGGCC  
AGTCTACAGG AAAGTTCTAC AACAGCAGAG GCCAGACAGT  
GCAACCCTTT CTGGGGGTGT  
GGGGGAGGAA AACCAACAA AACACAAAC TATAAAACAA  
AGAGAAGGCC GAGGACAAAG  
CTTAGCAATG CATACTTCCC TTTCTATGTG AAGCCCTGGG  
CTCCACCAGT ACTGCAGAAA  
GAAGCAAGCA ATGAGGGACA GGAGGTTGGC TCTAGGCCCA  
GGGGTTGTCA AAATAGTCCA  
CAGGCCAAAG GCAGCCTGAT GTCTGTTTTT ATAAACAAAA  
TTTTATTGGC ACACATTGGT  
TATGTATCAG CTAGGCTATT TTCATTACAA TAGAGGCCAT  
ATGGTCTGTA AAGTCTAAAA

TATTTACTCT GCTGTTTTAC ATAAAAAGTT GACAGACTCT  
TGCTCTAGAC TGACAAATAT  
CTAAGACCTT GTTTTCTGAG GTTCAAGTTT CAGAGGGGTC  
TCTGCAGCAA GTGGGTAAAG  
CTGGTCTAGG TCATGCTATG ATGTCTAGGG TCCCCTCAGA  
GTGGAAGGCC TGCTTAGCAC  
AAATGAAGTA AAGTAACTTG CTGGCTCTTT GTTCTTTTCT  
CCACACTCTA TACTTTAGCT  
CTGCCTC

18. Séquence d'ADN située en amont de la partie codante du gène de la souris de la protéine SM22 caractérisée en ce qu'elle comprend au moins une partie de la séquence SEQ ID n° 2 suivante, ou une séquence s'hybridant avec celle-ci dans des conditions stringentes.

## SEQ ID NO: 2 :

GGGCCCCAGG AATGTGTTTC CTTCTCTCCA CCATGTTTTT ATAGCTCTTG GGCTGGGAGA	
AGAGGCGGGT CTGGGTCTTT GTTCTGAGC TTTGTTCTAT GTTCCTCCAT GCTACGGTTG	1
CAATTGTTTT CTATGAACGA GTACATTCAA TAAAGACAAC CAGACCTGGG ATTTGGGGTC	1
TTACTGATGT GTTGGGAGGT GCAGGAGCCT CCGTGTCCCA TTTATTTTGG CCTTCCCGTC	2
TCGTTTCTGT GCGTGGCTAC ATTGGGAATG ACCTTCCTTG ATCCCACCAA GCCACCCATT	3
GATTCTGTAA ACATGTGACC CTTGCTCCAA GCATTGCTTA CAGGAGCAGG ATACTGAAAG	3
TGTGTCTGTG CCCTCTCCTG ATAACCCCTC CCTTCAGCAG GCACACAGCA CCTGACTACC	4
CACCACGTAT GTAAACGTCA GTATCCTTTC CAGCCAGCTC TGCAGATGGG TGTCCAGGCT	4
GTGCATGATG CACCTCAAGT GGGCAGAGCT TGCAGGCCAA GGTTTTAAAG GCTGTTTCAGG	5
AATGGATGGC AAGCAGGATC TAAGAGGAGG GGGGGTTGTT GTTGTTTGGG GGGGGGGTGG	6
TTTTGGTTTG TTTTTTTTGA GACAGGGTTT CTCTGTGTGG CCCTGGCCCT CCTGGAACCC	6
ACTCTGTAGA CCAGGCTGGC CTGGAACCTA GAAATCTGCC TGCCTCTGCC TCCCGAGTGC	7
TGGGATTAAA GGCCTGTGCC CATCGAGGAG GGAGATTTTA TTTAGATTAT AAAAAGGACG	7
GGATTTGGGG AATCCTGTCT AGTGAATTCA GGACGTAATC AGTGGCTGGG AAGCAAGAGC	8
TCTAGAGGAG CTCCAGCTTA TTATGACCCT TCCTTCAGAT GCCACAAGGA GGTGCTGGAG	9
TTCTATGCAC CAATAGCTTA AACCAGCCAG GCTGGCTGTA GTGGATTGAG CGTCTGAGGC	9
TGCACCTCTC TGGCCTGCAG CCAGTTCTGG GTGAGACTGA CCCTGCCTGA GGGTCTCTC	10
CCTCCCTCTC TCTACTCCTT CCTCCCTCTC CCTCTCCCTC TCTCTGTTTC CTGAGGTTTC	10



CAGAATTGGG	GATGGGACTC	AGAGACACCA	CTAAAGCCTT	ACCTTTTAAG	AAGTTGCATT	114
CAGTGAGTGT	GTGAGACATA	GCACAGATAG	GGGCAGAGGA	GAGCTGGTTC	TGTCTCCACT	120
GTGTTTGGTC	TTGGGTACTG	AACTCAGACC	ATCAGGTGTG	ATAGCAGTTG	TCTTTAACCC	126
TAACCCTGAG	CCTGTCTCAC	CTGTCCCTTC	CCAAGACCAC	TGAAGCTAGG	TGCAAGATAA	132
GTGGGGACCC	TTTCTGAGGT	GGTAGGATCT	TTCACGATAA	GGACTATTTT	GAAGGGAGGG	138
AGGGTGACAC	TGTCCTAGTC	CTCTTACCCT	AGTGTCTCCA	GCCTTGCCAG	GCCTTAAACA	144
TCCGCCCATT	GTCACCGCTC	TAGAAGGGGC	CACCCTTGAC	TTGCTGCTAA	ACAAGGCACT	150
CCCTAGAGAA	GATACCATAC	CTGTGGGCAG	GATGACCCAT	GTTCTGCCAT	GCACTTGGTA	156
GCCTTGGAAG	GGCCACTTTG	AACCTCAATT	TTCTCAACTG	TTAAATGGAG	TGGTAACTGC	162
TATCTCATAA	TAAAGGGGAA	CGTGAGGAAG	GCGTTTGGAT	AGTGCCTGGT	TGCGGCCAGG	168
CTGCAGTCAA	GACTAGTTCC	CACCAACTCG	ATTTTAAAGC	CTTGCAAGAA	GGTGGCTTGT	174
TTGTCCCTTG	CAGGTTCCCT	TGCTCGGGCC	AAACTCTAGA	ATGCCTCCCC	CTTTCTTTCT	180
CATTGAAGAG	CAGACCCAAG	TCCGGGTAAC	AAGGAAGGGT	TTCAGGGTCC	TGCCCCATAA	186
AGGTTTTTCC	CGGCCGCCCT	CAGCACCGCC	CCGCCCCGAC	CCCCGCAGCA	TCTCCAAAGC	192
ATGCAGAGAA	TGTCTCCGGC	TGCCCCCGAC	AGACTGCTCC	AACTTGGTGT	CTTTCCCCAA	198
ATATGGAGCC	TGTGTGGAGT	GAGTGGGGCG	GCCCCGGGTG	GTGAGCCAAG	CAGACTTCCA	204
TGGGCAGGGA	GGGGCGCCAC	GGGGCGGCAG	AGGGGTGACA	TCACTGCCTA	GGCGGCCTTT	210
AAACCCCTCA	CCCAGCCGGC	GCCCCGGCCC	GTCTGCCCCA	GCCCAGACAC	CGAAGCTACT	216
CTCCTTCCAG	TCCACAAACG	ACCAAGCCTT				219

19. Souche déposée le 25 mars 1996 auprès de la CNCM sous le n° I-1685 portant le plasmide p2126nlz comprenant la séquence selon la revendication 18.

20. Souche déposée le 25 mars 1996 auprès de la CNCM sous le n° I-1686 portant le plasmide p2126INTnlz comprenant la séquence selon la revendication 17.

21. Vecteur caractérisé en qu'il contient une séquence selon l'une des revendications 1 à 16.

22. Vecteur selon la revendication 20, caractérisé en ce qu'il contient une origine de répllication efficace dans les cellules des artères.

23. Vecteur selon l'une des revendications 21 et 22 caractérisé en ce qu'il est un dérivé d'un adénovirus.

24. ARN caractérisé en ce qu'il est capable d'être exprimé par une séquence ou un vecteur selon l'une des revendications 1 à 23.

25. Composition caractérisée en ce qu'elle contient une séquence ou un vecteur selon l'une des revendications 1 à 16 et 21 à 23.

26. Composition selon la revendication 25, caractérisée en ce que la séquence ou le vecteur sont compris dans une composition facilitant leur transfection dans les cellules.

27. Composition selon la revendication 25, caractérisée en ce qu'elle comprend un gel qui est un complexe de poly-L-lysine et de lactose.

28. Médicament caractérisé en ce qu'il contient une séquence ou un vecteur selon l'une des revendications 1 à 16 et 21 à 23.

29. Composition pharmaceutique caractérisée en ce qu'elle contient une quantité pharmaceutiquement efficace d'une séquence nucléique ou d'un vecteur selon l'une des revendications 1 à 16 et 21 à 23, et des excipients pharmaceutiquement compatibles.

30. Utilisation d'un vecteur ou d'une séquence selon l'une des revendications 1 à 16 et 21 à 23 pour la fabrication d'un médicament pour le traitement des maladies coronariennes.

31. Utilisation d'un vecteur ou d'une séquence selon l'une des revendications 1 à 16 et 21 à 23 pour la fabrication d'un médicament pour le traitement de la resténose.

32. Utilisation d'un vecteur ou d'une séquence selon l'une des revendications 1 à 16 et 21 à 23 pour la fabrication d'un médicament pour le traitement des mutations fragilisant les vaisseaux.

5 33. Animaux transgéniques caractérisés en ce qu'ils portent une séquence ou un vecteur selon l'une des revendications 1 à 16 et 21 à 23, dans lequel le gène codant pour la protéine d'intérêt thérapeutique est remplacé par un gène reporteur.

10 34. Procédé de criblage de molécules in vitro pour leur activité sur les séquences régulatrices du gène codant pour la protéine SM22 comprenant les étapes suivantes:

- transfection de cellules avec une séquence ou un vecteur selon l'une des revendications 1 à 16 et 21 à 23, dans lequel le gène codant pour la protéine d'intérêt thérapeutique est remplacé par un gène reporteur,

15 - incubation des cellules transfectées avec la molécule à tester, et  
- quantification de l'expression du gène réporteur.

20 35. Procédé de détection de mutations sur la région comportant la séquence selon l'une des revendications 1 à 16 ou le vecteur selon l'une des revendications 21 à 23, caractérisé par une altération de l'expression du gène placé en aval de ladite séquence.

36. Procédé d'expression de protéines d'intérêt thérapeutique caractérisé en ce qu'il met en oeuvre les produits selon l'une des revendications 1 à 24.

1/25

Fig. 1

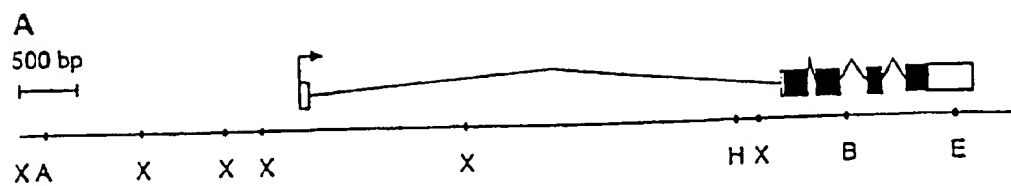


FIG. 1

Rep	-1506	1	AGACAGGGT	TCTCTCTGTGGCCCTGGCCCTCC	TGGAAACCCAC	TCTGT-AGACCAGGCTG	60
Rep	702	1	.....	.....A.....TG.T.....T.T.....	.....	.....	60
Rep	1830	1	.....	.....A.A.....A..TA.....T.....-	.....	.....	60

61	GCC	TGAA	-CTCAGAAATCTG	CCCTGCCCTCTGCCCTCCCGAGT	GCTGGGATTAAGCGGTGCC	123
61	....	C...T.....	.....A.....	.....	.....	123
61	....	CC..-.....	.....	.....C.....	.....C...	123

FIG. 2

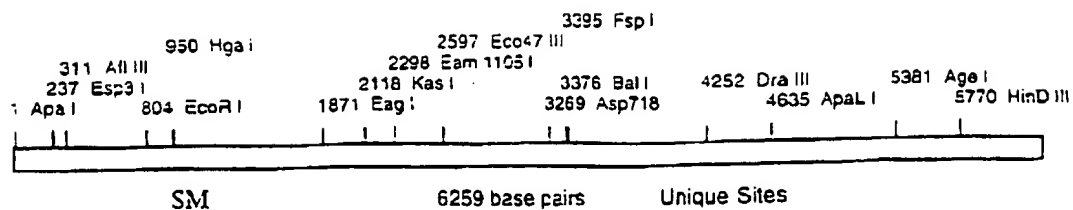
3/25



FIG. 3

Fig 4 A

5/25

Fig 4B



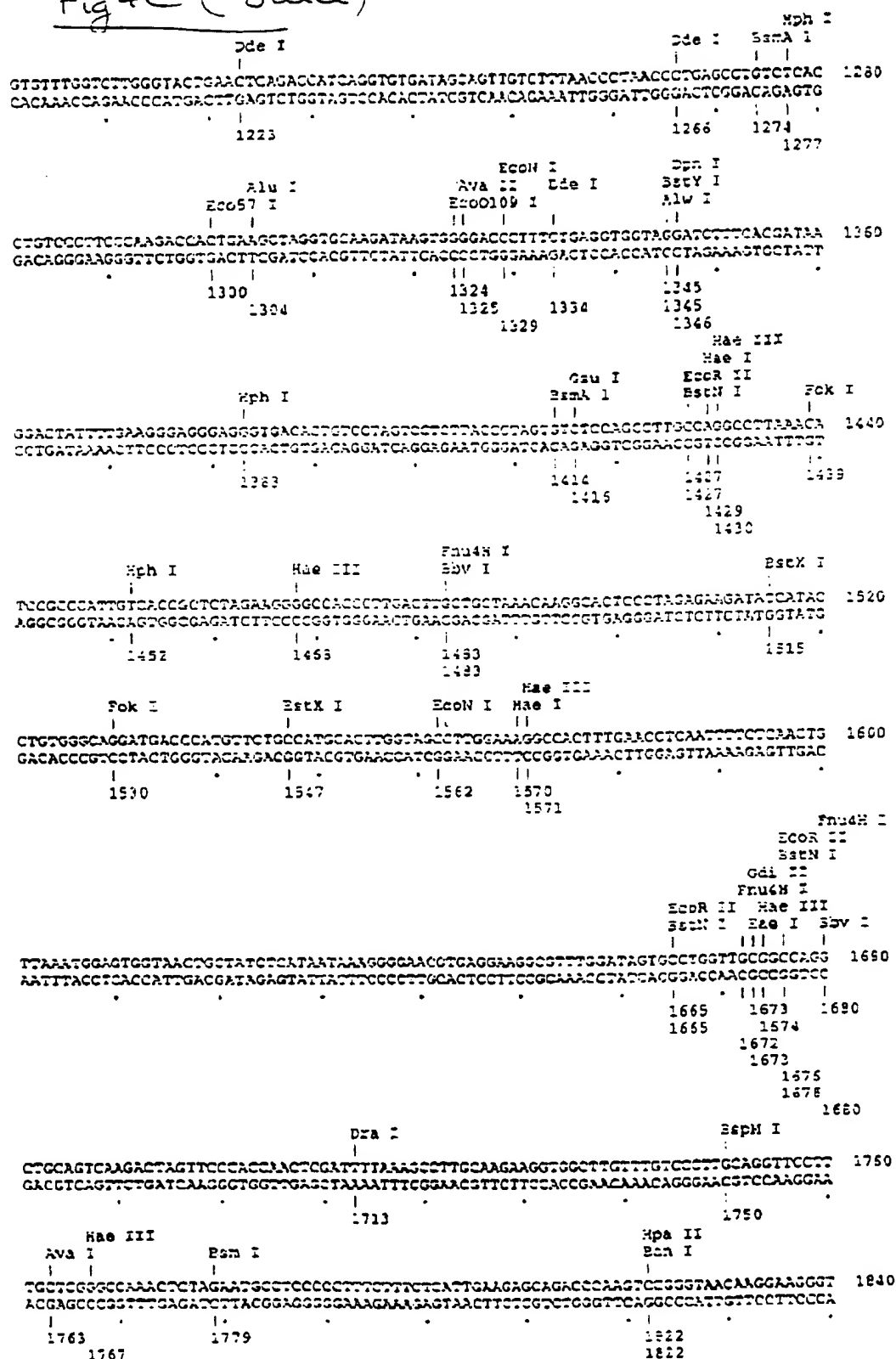
[illegible]

Fig 4C (Suite)

[illegible]

**8/25**

Fig 4C (Suite)



**9/25**

Fig 4C (Suite)

Fnu4H I  
 Hae III  
 Ggi II  
 Eae I  
 Hpa II  
 Bcn I Dde I  
 Fnu4H I  
 Bbv I  
 Ava II  
 EcoO109 I  
 1345  
 1346  
 1869  
 1870  
 1871  
 1871  
 1871  
 1872  
 1873  
 Fnu4H I  
 Bbv I  
 Hpa II  
 BsmA I  
 1932  
 1936  
 1939  
 1939  
 Hpa II  
 Bcn I  
 Bcn I  
 Ava I  
 Hae III  
 Fnu4H I Hph I  
 2008 2020 2053 2054 2074  
 2010  
 2012  
 2012  
 2013  
 2013  
 Bcn I  
 Hinf I  
 Hha I  
 Hae I  
 Hae II  
 Bba I  
 Ban I  
 Aha II  
 Fnu4H I Hph I  
 2053 2054 2054  
 Bcn I  
 Hinf I  
 Hha I  
 Hae I  
 Hae II  
 Bba I  
 Ban I  
 Aha II  
 Fnu4H I Dfa I  
 Bgl I Hae III Hph I Cfr10 I Hpa II Hae III  
 2086 2094 2095 2108 2115 2124 2154  
 2087 2092 2116 2118 2118 2118 2119 2119 2123  
 Bsp1286 I Dpn I  
 Ban II BstY I  
 2223 2233 2234 2234  
 CTCCTTCCAGTCCACAAACGACCAAGCCCTGTATAGTCCAAATCATGCGGAGAGAGGGCGTGTGGGCTCAATTAGATCCCC 2240  
 CAGGAGGCTCAGGTGTTTGGCTCGGAAACATTCACGTTTCAGTACCGCTCGTCTTCCCGACACCCGAGTTAATCTAGGGG

Fig 4C (Suite)

[illegible]

11/25

Fig 4C (Suite)

[illegible]

Fig 4C (Suite)

[illegible]

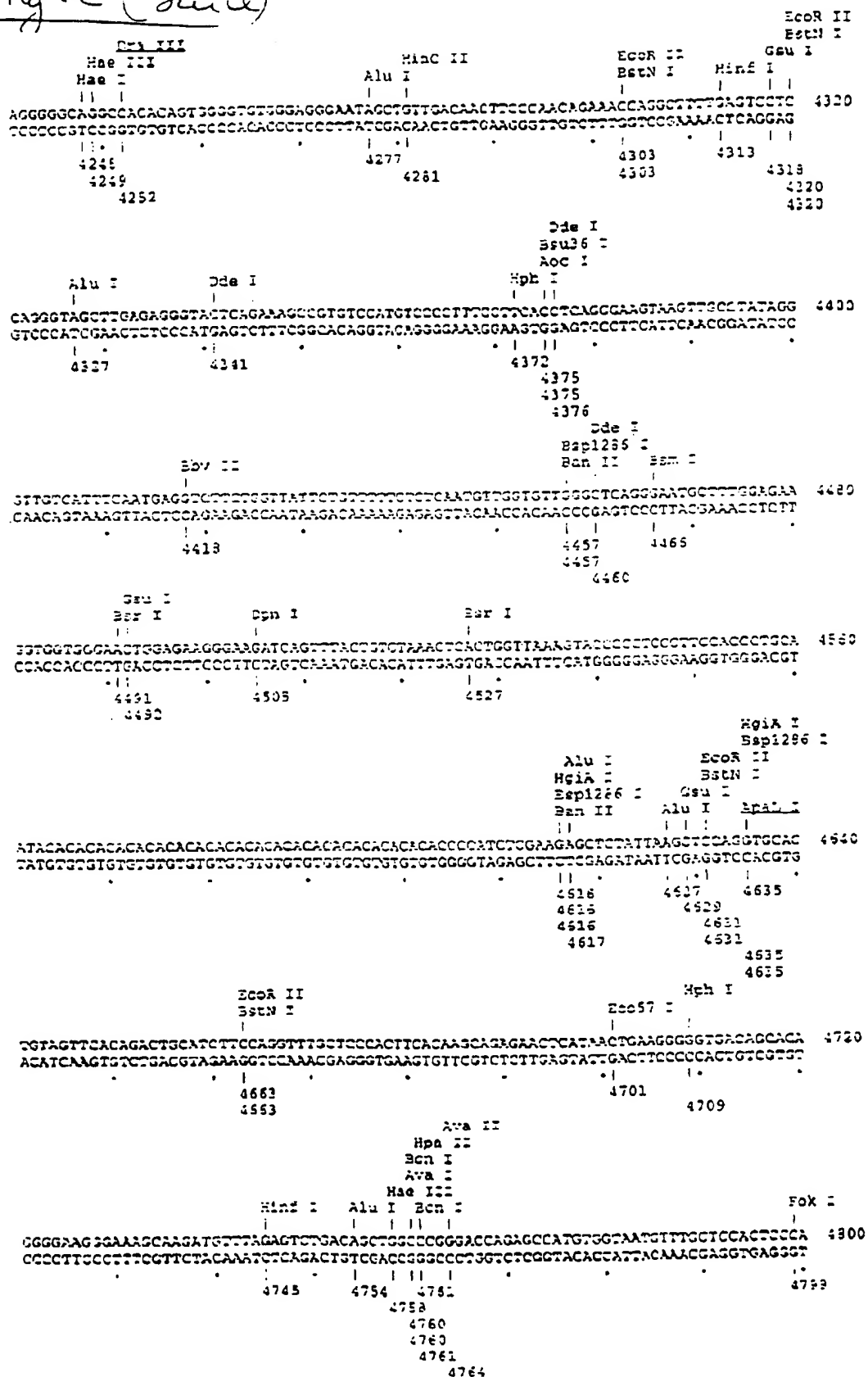
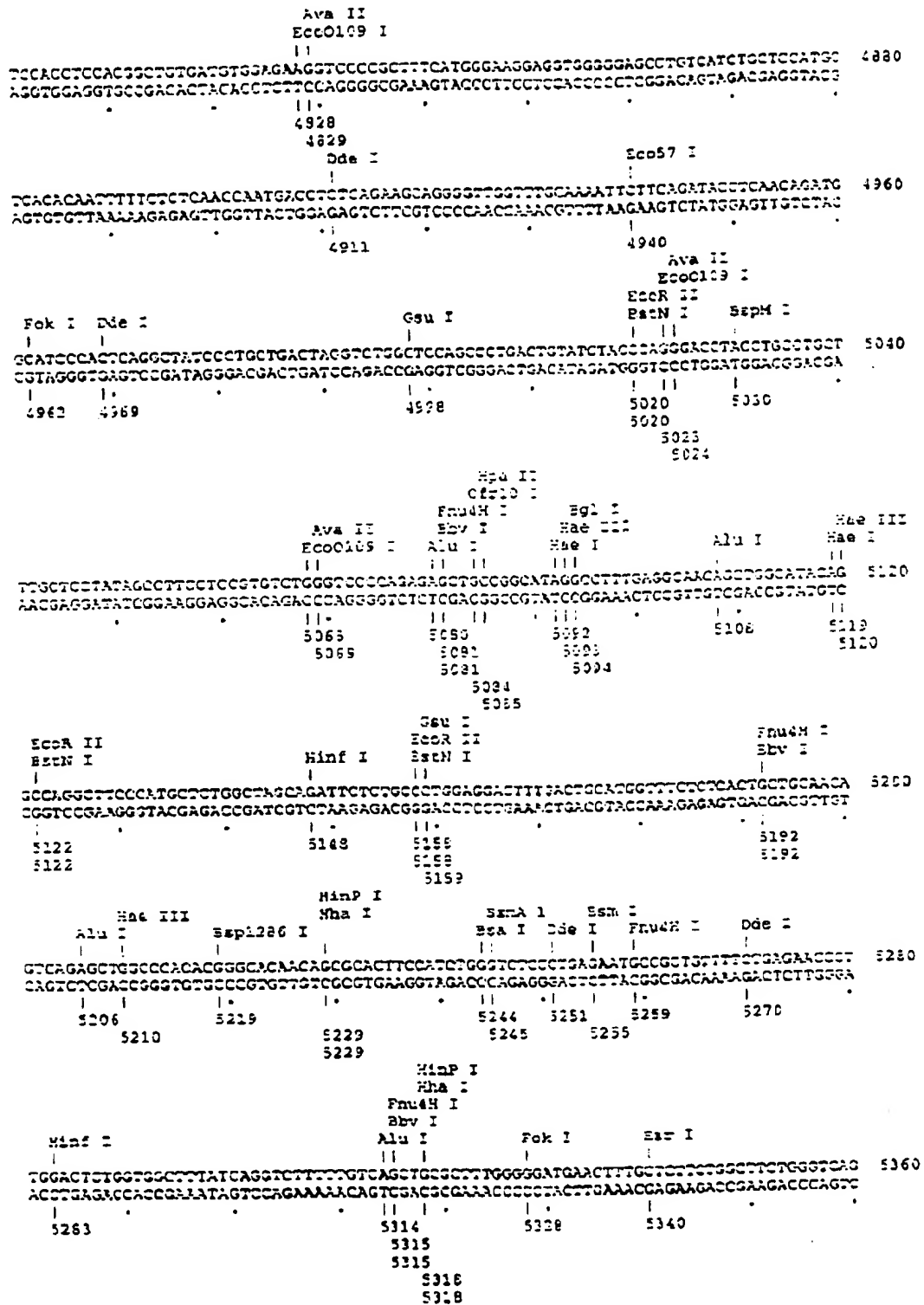




Fig 4C (Suite)

14/25



15/25

[illegible]

Fig 4C (Suite)

TATATAACAAATTTTATTGGACACATGGGTTATGTATCAGCTAGGCTATTTTCATTACAAATAGAGGGCCATATGGTCTG 6000  
 AATATTTTCTTTAAATAACCGTGTGTACCCATATACATAGTCGATCCGATAAAGTAATGTTATCTCCGGTATACCGAGC  
 5951 5986 5997  
 TAAAGTCTAAATATTTACTCTGCTGTTTACATAAAAGTTGACAGACTCTTGGCTCTAGAGTGACAAATATCTTAGAGCC 6080  
 ATTTTCAGATTTTTAAATGAGACGACAAATGTATTTTCAACTGCTCTGAGAGCGAGATCTGACTGTTTATAGATTCTGG  
 6040 6047 6073  
 TTGTTTCTGAGGTTTCAAGGTTCTAGAGGGGCTCTCTGACAGCAAGTGGGTAAGCTGGTCTAGGCTATGCTATGATGCTCTAG 5160  
 AACAAATAGACTTCCAGTTTCAAGTCTCTCCCGACAGACTCTGTTTCAAGCTTTTCGACCCAGATCCGCTATCAGTACTACAGATC  
 6088 6109 6116 6131 6160  
 6110 6116  
 GGTCCCTTCAGAGTGGGAGGGCTGCTTTAGCACAATCGAATTAAGTTAACTTGGTGGCTCTTTTCTTTTCTCCAGCTC 6240  
 CCAGCGGAGTCTCAGCTTCCGGACGAATCTGTGTTTACTTCATTTCTTACAGCCGAGAAACAAGAAAGAGGCTCTGAG  
 6151 6172 6183  
 6167 6175  
 6194 6194  
 TATACTTTAGGCTCTGGCTC 6259  
 ATATGAAATCCAGACGGAG  
 6245

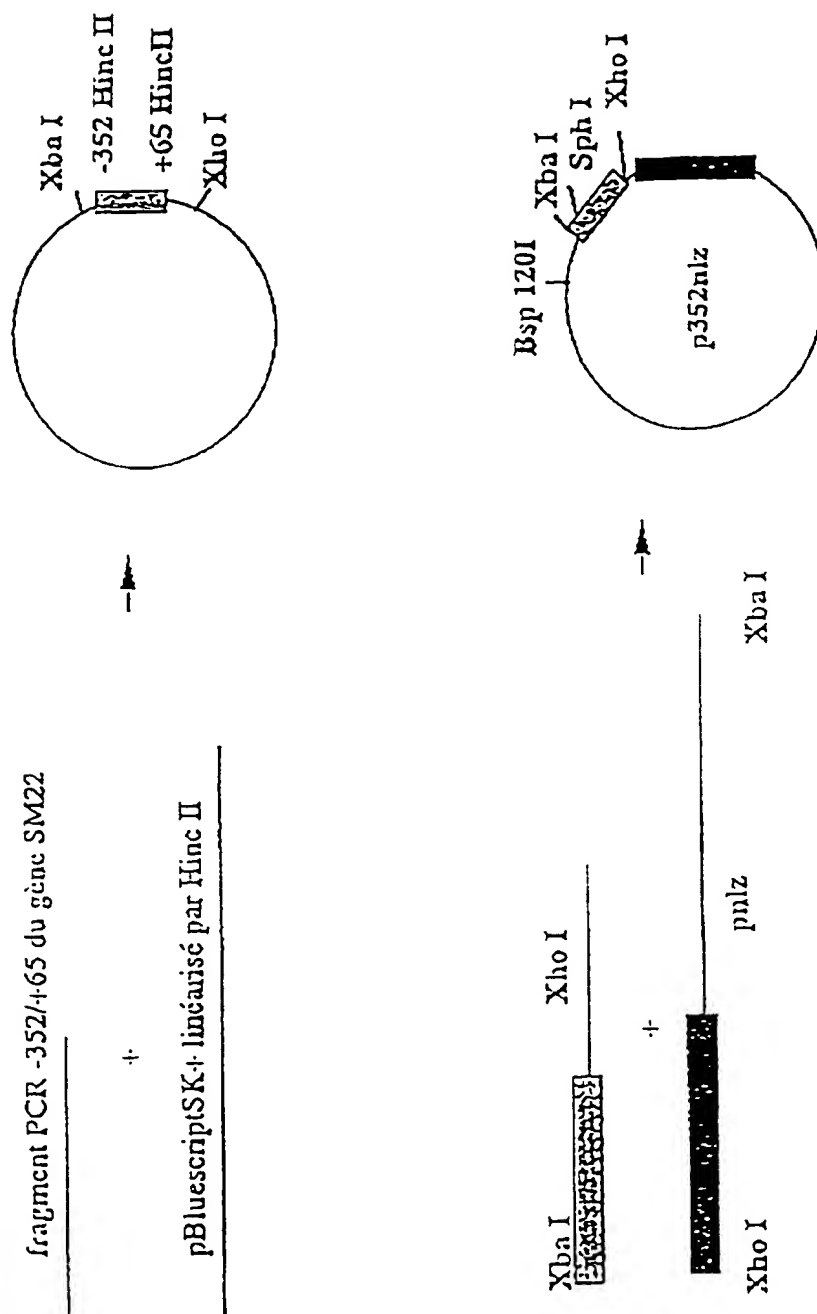


FIG - 5

18/25

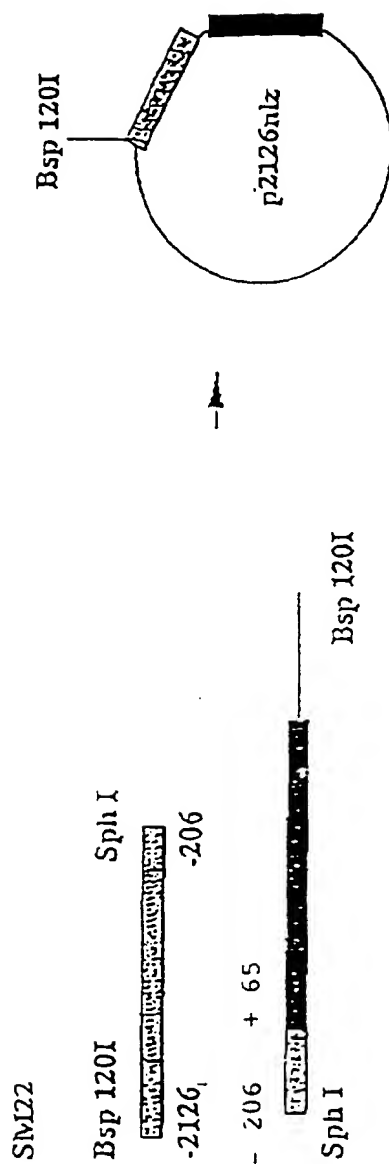


FIG - 6

19/25

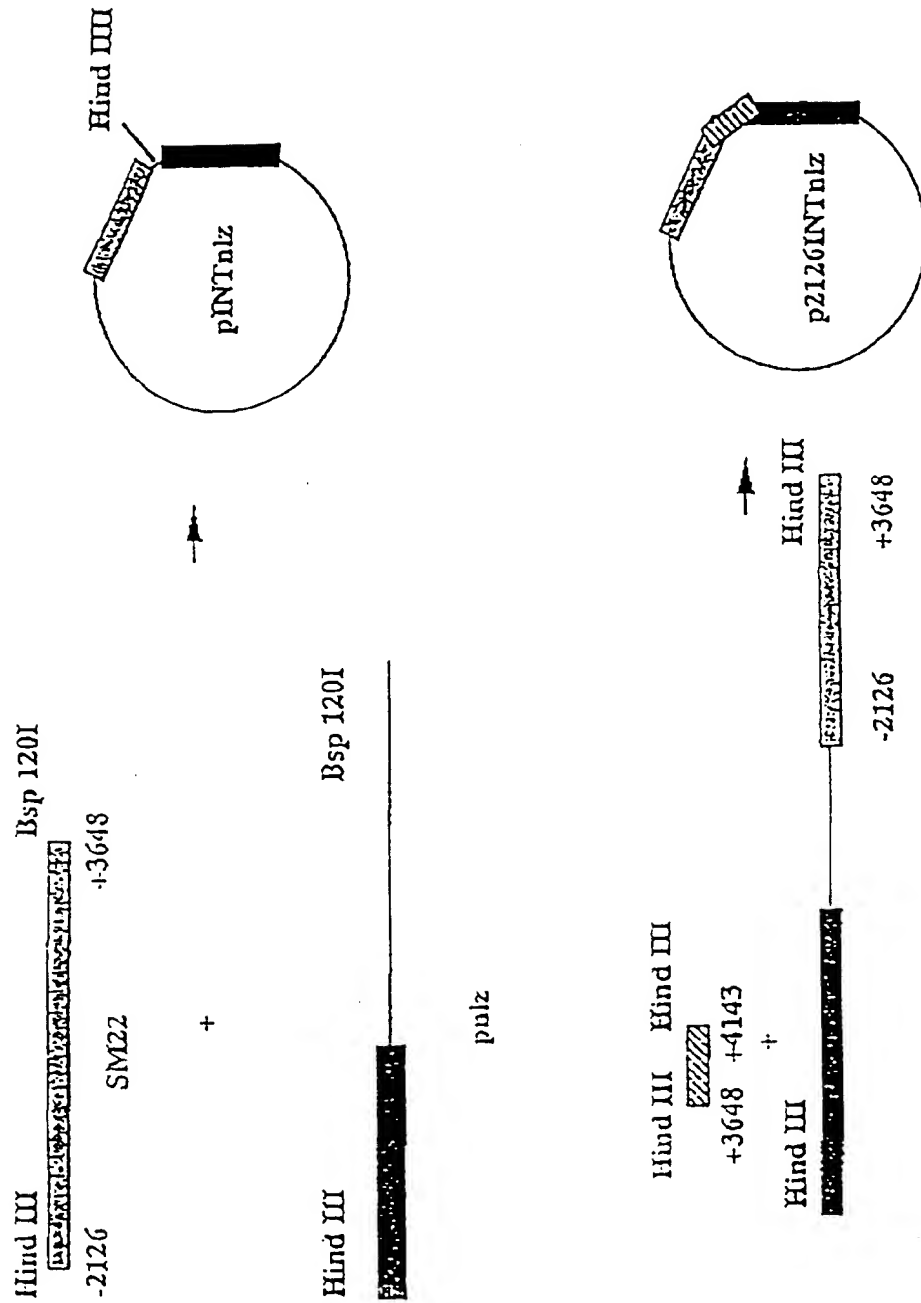


FIG - 7



A

FIG. 8A



B

FIG. 8B

21/25

FIG. 9A



FIG. 9B

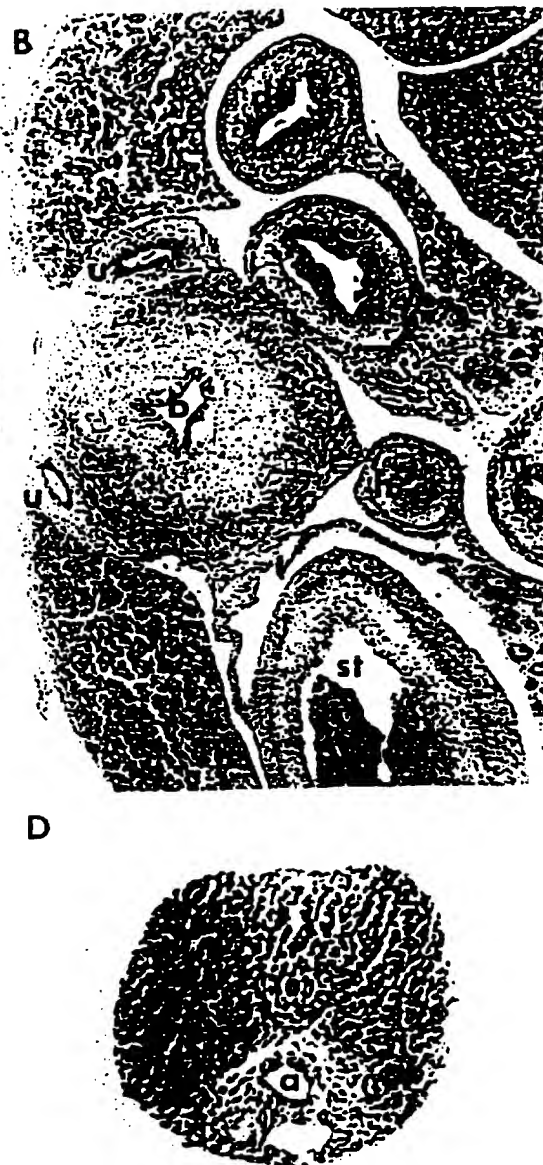


FIG. 9C

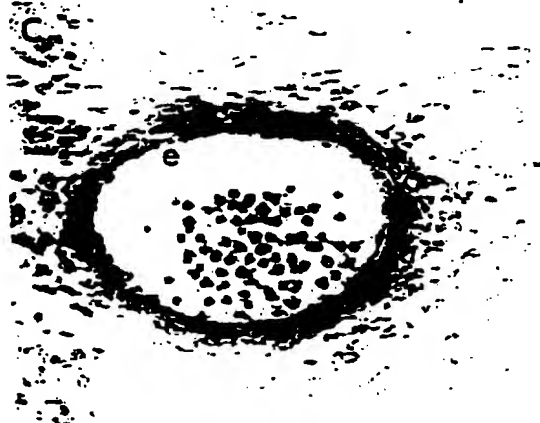


FIG. 9D





FIG. 10A

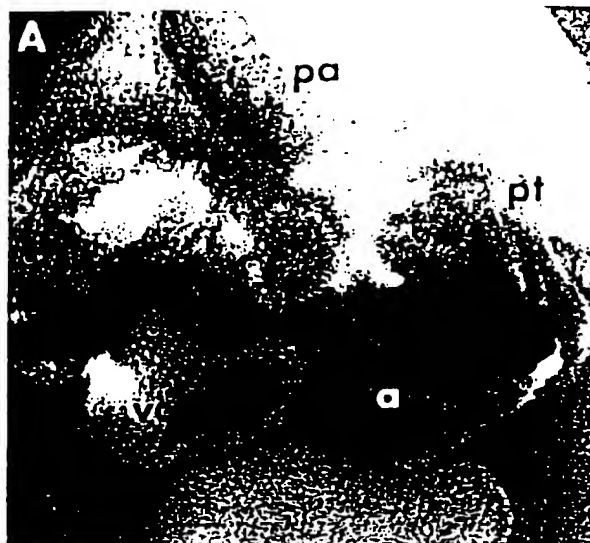
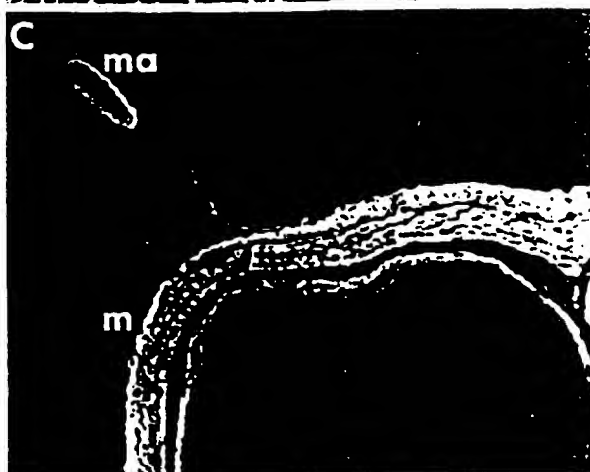


FIG. 10B



FIG. 10C



23/25

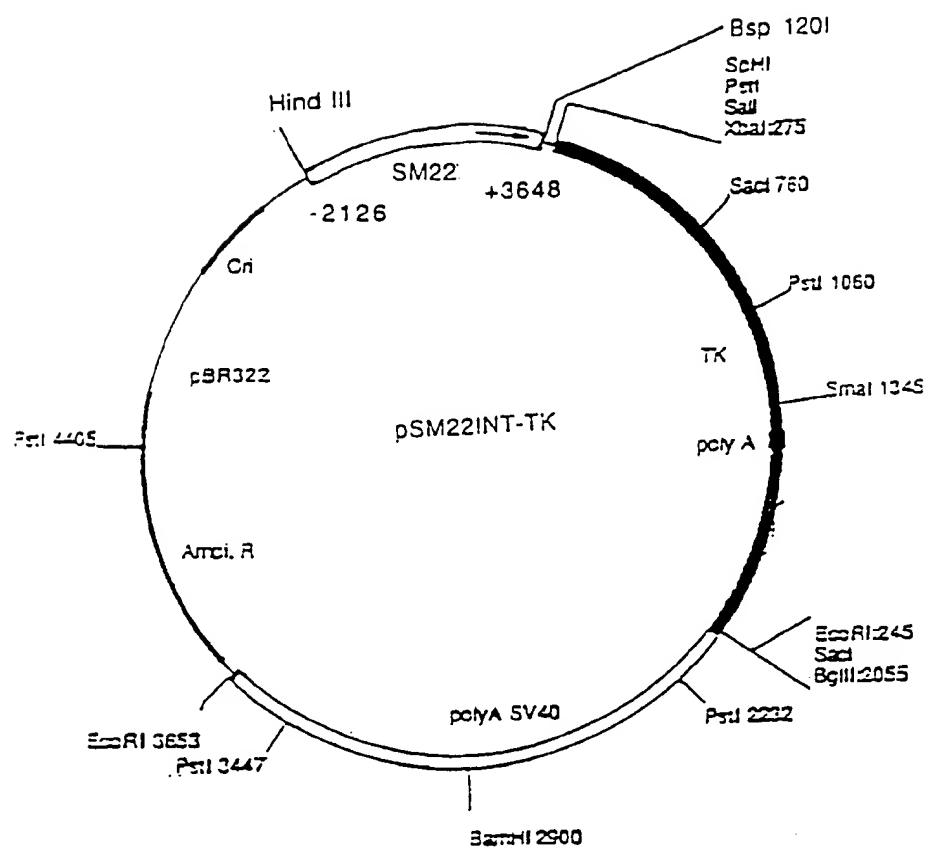


FIG. 11

Fig 12 B



Fig 12 A



Fig 12D



Fig 12C



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 97/00543

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12N15/12 C07K14/47 A61K31/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K A61K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	THE JOURNAL OF CELL BIOLOGY, vol. 132, 1996, pages 849-859, XP000609076 L. LI ET AL: "Expression of the SM22 alpha promoter in transgenic mice provides evidence for distinct transcriptional regulatory programs in vascular and visceral smooth muscle cells" cited in the application see page 850 : results, characterization of the SM22 alpha gene*	15,17, 32,33
A	---	1-14
	-/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

5 August 1997

Date of mailing of the international search report

26. 08. 97

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (- 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (- 31-70) 340-3016

Authorized officer

Mateo Rosell, A.M.

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/FR 97/00543

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 270, no. 22, 1 June 1995, USA, pages 13460-13469, XP000609071 J. SOLWAY ET AL.: "Structure and expression of a smooth muscle cell-specific gene, SM22 alpha" cited in the application see figure 18 ---	15,17
X	BIOCHEMICAL JOURNAL, vol. 310, 1995, GREAT BRITAIN, pages 1037-1043, XP000196609 P.R. KEMP ET AL.: "cloning and analysis of the promoter region of the rat SM22 alpha gene" cited in the application see figure 4 ---	15,17
A	WO 94 13824 A (UNIVERSITE PIERRE ET MARIE CURIE) 23 June 1994 see page 1, paragraph 1 see page 3, paragraph 6 - page 4, paragraph 2 ---	1-7
A	DATABASE WPI Section Ch, Week 9141 Derwent Publications Ltd., London, GB; AN 91-300978 XP002020314 ODA JUNICHI: "New DNA sequence of gene coding lipoprotein lipase-derived from microbe, pref. belonging to Pseudomonas genus" & JP 03 201 989 A ((TOYM) TOYOBO KK) , 3 September 1991 see abstract ---	1-5,10
A	WO 95 21258 A (UNITED STATES OF AMERICA, DEPARTMENT OF HEALTH AND HUMAN SERVICES) 10 August 1995 see page 9, line 12-34 see page 22, line 20-36 see claims 1,7,8 ---	1-5,11
A	EP 0 174 608 A (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 19 March 1986 see abstract ---	1-5,13
A	WO 95 09916 A (RHONE POULENC RORER SA) 13 April 1995 see page 2, line 13 - page 3, line 2; claims 9,11,14,15 ---	1-5,14
	-/--	

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/FR 97/00543

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 95 19167 A (MALLINCKRODT MEDICAL INC) 20 July 1995 see page 3, line 18 - line 34 ---	1-5,8, 11,14
A	SCIENCE, vol. 267, 27 January 1995, pages 518-522, XP000609101 M.W. CHANG ET AL.: "Cytostatic gene therapy for vascular proliferative disorders with a constitutively active form of the retinoblastoma gene product" cited in the application see the whole document ---	8,21,22
A	WO 95 07361 A (INSTITUT PASTEUR) 16 March 1995 see the whole document	34
A	& FR 2 709 761 A (PASTEUR INSTITUT) 17 March 1995 cited in the application ---	
A	NUCLEIC ACIDS RESEARCH, vol. 21, no. 4, 25 February 1993, GB, pages 871-878, XP000371764 P. MIDOUX ET AL.: "Specific gene transfer mediated by lactosylated poly-L-lysine into hepatoma cells" cited in the application see the whole document ---	26
P,X	DEVELOPMENT, vol. 122, 1 August 1996, GREAT BRITAIN, pages 2415-2425, XP000196601 H. MOESSLER ET AL.: "The SM 22 promoter directs tissue-specific expression in arterial but not in venous or visceral smooth muscle cells in transgenic mice" see figure 2 -----	15,17, 32,33

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 97/00543

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9413824 A	23-06-94	FR 2699191 A CA 2150536 A EP 0674716 A JP 8506722 T	17-06-94 23-06-94 04-10-95 23-07-96
WO 9521258 A	10-08-95	AU 1690295 A CA 2182498 A EP 0742830 A	21-08-95 10-08-95 20-11-96
EP 0174608 A	19-03-86	JP 61108389 A	27-05-86
WO 9509916 A	13-04-95	FR 2710846 A AU 7816294 A CA 2173338 A EP 0722496 A FI 961494 A NO 961220 A	14-04-95 01-05-95 13-04-95 24-07-96 03-04-96 26-03-96
WO 9519167 A	20-07-95	CA 2181035 A	20-07-95
WO 9507361 A	16-03-95	FR 2709761 A CA 2171469 A EP 0717781 A	17-03-95 16-03-95 26-06-96



# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No  
PCT/FR 97/00543

**A. CLASSEMENT DE L'OBJET DE LA DEMANDE**  
CIB 6 C12N15/12 C07K14/47 A61K31/00

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

**B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE**

Documentation minimale consultée (système de classification suivi des symboles de classement)  
CIB 6 C07K A61K C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)

**C. DOCUMENTS CONSIDERES COMME PERTINENTS**

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	THE JOURNAL OF CELL BIOLOGY, vol. 132, 1996, pages 849-859, XP000609076 L. LI ET AL: "Expression of the SM22 alpha promoter in transgenic mice provides evidence for distinct transcriptional regulatory programs in vascular and visceral smooth muscle cells" cité dans la demande *voir page 850 : results, characterization of the SM22 alpha gene*	15,17, 32,33
A	--- -/-	1-14

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

\* Catégories spéciales de documents cités:

- "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

- "T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
- "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
- "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
- "&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

5 Août 1997

Date d'expédition du présent rapport de recherche internationale

26. 08. 97

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tél. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+ 31-70) 340-3016

Fonctionnaire autorisé

Mateo Rosell, A.M.

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No  
PCT/FR 97/00543

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Categorie *	Identification des documents cites, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 270, no. 22, 1 Juin 1995, USA, pages 13460-13469, XP000609071 J. SOLWAY ET AL.: "Structure and expression of a smooth muscle cell-specific gene, SM22 alpha" cité dans la demande voir figure 18</p> <p style="text-align: center;">---</p>	15,17
X	<p>BIOCHEMICAL JOURNAL, vol. 310, 1995, GREAT BRITAIN, pages 1037-1043, XP000196609 P.R. KEMP ET AL.: "cloning and analysis of the promoter region of the rat SM22 alpha gene" cité dans la demande voir figure 4</p> <p style="text-align: center;">---</p>	15,17
A	<p>WO 94 13824 A (UNIVERSITE PIERRE ET MARIE CURIE) 23 Juin 1994 voir page 1, alinéa 1 voir page 3, alinéa 6 - page 4, alinéa 2</p> <p style="text-align: center;">---</p>	1-7
A	<p>DATABASE WPI Section Ch, Week 9141 Derwent Publications Ltd., London, GB; AN 91-300978 XP002020314 ODA JUNICHI: "New DNA sequence of gene coding lipoprotein lipase-derived from microbe, pref. belonging to Pseudomonas genus" &amp; JP 03 201 989 A ((TOYM) TOYOBO KK) , 3 Septembre 1991 voir abrégé</p> <p style="text-align: center;">---</p>	1-5,10
A	<p>WO 95 21258 A (UNITED STATES OF AMERICA, DEPARTMENT OF HEALTH AND HUMAN SERVICES) 10 Août 1995 voir page 9, ligne 12-34 voir page 22, ligne 20-36 voir revendications 1,7,8</p> <p style="text-align: center;">---</p>	1-5,11
A	<p>EP 0 174 608 A (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 19 Mars 1986 voir abrégé</p> <p style="text-align: center;">---</p>	1-5,13
A	<p>WO 95 09916 A (RHONE POULENC RORER SA) 13 Avril 1995 voir page 2, ligne 13 - page 3, ligne 2; revendications 9,11,14,15</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	1-5,14

1

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No  
PCT/FR 97/00543

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO 95 19167 A (MALLINCKRODT MEDICAL INC) 20 Juillet 1995 voir page 3, ligne 18 - ligne 34 ---	1-5,8, 11,14
A	SCIENCE, vol. 267, 27 Janvier 1995, pages 518-522, XP000609101 M.W. CHANG ET AL.: "Cytostatic gene therapy for vascular proliferative disorders with a constitutively active form of the retinoblastoma gene product" cité dans la demande voir le document en entier ---	8,21,22
A	WO 95 07361 A (INSTITUT PASTEUR) 16 Mars 1995 voir le document en entier	34
A	& FR 2 709 761 A (PASTEUR INSTITUT) 17 Mars 1995 cité dans la demande ---	
A	NUCLEIC ACIDS RESEARCH, vol. 21, no. 4, 25 Février 1993, GB, pages 871-878, XP000371764 P. MIDOUX ET AL.: "Specific gene transfer mediated by lactosylated poly-L-lysine into hepatoma cells" cité dans la demande voir le document en entier ---	26
P,X	DEVELOPMENT, vol. 122, 1 Août 1996, GREAT BRITAIN, pages 2415-2425, XP000196601 H. MOESSLER ET AL.: "The SM 22 promoter directs tissue-specific expression in arterial but not in venous or visceral smooth muscle cells in transgenic mice" voir figure 2 -----	15,17, 32,33

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No

PCT/FR 97/00543

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9413824 A	23-06-94	FR 2699191 A	17-06-94
		CA 2150536 A	23-06-94
		EP 0674716 A	04-10-95
		JP 8506722 T	23-07-96
-----			
WO 9521258 A	10-08-95	AU 1690295 A	21-08-95
		CA 2182498 A	10-08-95
		EP 0742830 A	20-11-96
-----			
EP 0174608 A	19-03-86	JP 61108389 A	27-05-86
-----			
WO 9509916 A	13-04-95	FR 2710846 A	14-04-95
		AU 7816294 A	01-05-95
		CA 2173338 A	13-04-95
		EP 0722496 A	24-07-96
		FI 961494 A	03-04-96
		NO 961220 A	26-03-96
-----			
WO 9519167 A	20-07-95	CA 2181035 A	20-07-95
-----			
WO 9507361 A	16-03-95	FR 2709761 A	17-03-95
		CA 2171469 A	16-03-95
		EP 0717781 A	26-06-96
-----			